

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
11 April 2002 (11.04.2002)

PCT

(10) International Publication Number  
**WO 02/29068 A2**

- (51) International Patent Classification<sup>7</sup>: **C12N 15/82**, 15/63, 15/12
- (74) Agent: **WÄCHTERSCHÄUSER, Günter**; Wächtershäuser & Hartz, Patent Attorneys, Tal 29, 80331 München (DE).
- (21) International Application Number: **PCT/EP01/11629**
- (22) International Filing Date: 8 October 2001 (08.10.2001)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
100 49 587.7 6 October 2000 (06.10.2000) DE
- (71) Applicant (for all designated States except US): **ICON GENETICS AG** [DE/DE]; Maximilianstr. 38/40, 80539 München (DE).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): **GLEBA, Yuri** [UA/US]; Muehlweg 37, 06114 Halle (DE). **DOROKOV, Yurii** [RU/RU]; rofsojuznaja Street, 146-3-187, Moskau 117321 (RU). **IVANOV, Peter** [RU/RU]; Vavilova Street, 37a-18, Moscow, 117312 (RU). **ATABEKOV, Joseph** [RU/RU]; Lomonosovski prospect, 15-142, Moscow 117311 (RU).
- (81) Designated States (*national*): AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:  
— without international search report and to be republished upon receipt of that report
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: VECTOR SYSTEM FOR PLANTS

(57) Abstract: The invention describes virus-based amplification vectors for plants containing additional plant-specific internal ribosome entry site (IRES) element(s) allowing for a polycistronic translation and a cap-independent translation of : a) heterologous gene(s); b) whole viral genome or c) viral subgenomic RNAs. Said IRES elements are of plant viral origin, or they are isolated from other organisms or engineered using different synthesis procedures. Said IRES element(s) and said heterologous gene(s) are inserted into amplification vectors and allow for the expression of said heterologous gene(s) in the absence of additional viral promoters, in particular, said expression is achieved through cap-independent translation.

**BEST AVAILABLE COPY**

## Vector system for plants

### FIELD OF INVENTION

This invention relates to a vector capable of amplification and expression and/or suppression of a gene in a plant, as well as uses thereof, notably use for producing a protein, and a method and pro-vector for generating said vector. The invention further relates to a gene expression system, notably for the expression of a pharmaceutical protein or for the expression of two or more genes in the same plant or plant cell.

### BACKGROUND OF THE INVENTION

Vectors for genetic engineering of plants are highly desirable for the production of proteins, for endowing a host plant with a new trait, for suppressing a gene of the host plant, or for determining the function of a gene, notably a gene determined by genomics.

Vectors, notably viral vectors, for the genetic engineering of plants are already known. These must be capable of infection, amplification and movement (both cell-to-cell and long-distance) in a plant in addition to having at least one sequence for gene expression or suppression. Prior art vectors rely on subgenomic promoters as transcriptional elements. A subgenomic promoter has the effect that, in a transfected plant cell, transcription of a vector nucleic acid sequence starts in part at said subgenomic promoter to generate a shorter RNA so that translation of a gene downstream from said promoters by the plant translation machinery is enabled. Translation may then proceed cap-dependent. Such multiple transcriptions are kinetically disadvantageous because of waste of replicase capacity.

Such vectors have a number of further shortcomings. The introduction of a virus subgenomic promoter into a vector sequence makes said sequence longer and thus less efficient. Moreover, the presence of several identical or similar subgenomic promoters which are well adapted to transcription in the host gives rise to frequent recombination events and instability with loss of sequence portions. On the other hand, if significantly different subgenomic promoters are used, recombination may be suppressed but such promoters may be too different to be effectively recognized by the transcription system, which means loss of efficiency. Moreover, vectors are usually highly integrated entities with several interdependent functional elements or genes tightly packed into a sequence. This is the reason why the operability of a vector for certain heterologous genes or the like is somewhat idiosyncratic and frequently gives unpredictable results, notably in terms of infectivity and expression. Further, the available sequence space for promoters is usually constrained if sequence overlaps with upstream genes

are present.

Therefore, it is an object of this invention to provide a novel vector for plant genetic engineering which is capable of efficient and stable operation in a host plant. It is a further object to provide a vector which is capable of high-level expression of a gene in a plant.

It has been surprisingly found that these objects can be achieved with a vector capable of amplification and expression of a gene in a plant comprising a nucleic acid having a sequence for at least one non-viral gene to be expressed and having or coding for at least one IRES element necessary for translation of a gene downstream thereof.

It has been previously suggested (WO 98/54342) to use a plant IRES element in a recombinant DNA molecule that has merely the function of gene expression (after integration into the host genome). However, the expression level is low. The exact reasons for this low expression level are not known. In any event, expression is limited to the very plant cells transformed, thus the overall efficiency in whole plants is extremely low.

It has been surprisingly found that it is possible to construct a plant vector which, when introduced into a plant cell, has not only the capability of gene expression but which has several additional functions which are all required for amplification and spreading throughout the plant so that the overall efficiency is extremely high. These functions comprise infection, amplification, cell-to-cell movement and long-distance movement. It is surprising that the required high degree of integration of functional and structural elements on a vector does not impair gene expression from said vector.

The IRES element of said vector can be located upstream of said non-viral gene to be expressed for directly supporting its translation. Alternatively, said IRES element may indirectly support the translation of said gene to be expressed by directly supporting the translation of another gene essential for a function of said vector selected from the group of infection, amplification, virus assembly, ability to suppress the silencing of viral infection development in plant cells, ability to redirect the metabolism in plant cells, and cell-to-cell or long-distance movement of said vector.

Further said vector may comprise at least a portion of a sequence of the host plant genome in an anti-sense orientation for suppressing a gene of the host plant.

It is a further object to provide a vector which is capable of the effective suppression of a gene in a plant. This object has been achieved by a vector capable of amplification in a plant

comprising a nucleic acid having or coding for at least one IRES element necessary for translation of a gene required for amplification of said vector and located downstream of said IRES element, said vector further comprising at least a portion of a sequence of the host plant genome in an anti-sense orientation for suppressing a gene of the host plant.

Further preferred embodiments are defined in the subclaims.

Here, the first plant expression and amplification vectors based on plant active translational (IRES) elements are described. Existing IRES elements isolated from animal viruses do not support translation in plant cells. Therefore, knowledge accumulated in animal expression systems is not applicable to plants. Animal IRES elements have never been tested for other functional properties, such as residual promoter activity, so this invention discloses the first bona fide cases of gene expression in plants relying exclusively on translation rather than on transcription with a subgenomic promoter necessary for expression of a gene downstream thereof.

The vectors of this invention allows preferably for regulation and preferential expression of a gene of interest in a plant by suppressing cap-dependent translation. In another preferred embodiment, very short homologous or artificial IRES elements are used, thus adding to the stability of the resulting vectors.

A preferred advantage of this strategy is that IRES sequences can be inserted upstream or downstream of viral gene(s) (e.g. the coat protein gene of tobacco mosaic virus) such that translation of downstream foreign gene(s) or the viral gene(s), respectively, may occur via a cap-independent internal ribosome entry pathway. Thus, said cap-independent translation of foreign gene(s) will occur from bicistronic or/and polycistronic RNAs.

#### General Problem Situation and Definitions

Upon infection of a plant with a virus, the early events of viral infection (entry and genome uncoating) occur. Then the virus must engage in activities that enable its genome to be expressed and replicated. The viral genome may consist of one (monopartite) or more (multipartite) RNA or DNA segments, and each of these segments may under certain conditions be capable of replicating in the infected cell. A viral replicon has been defined as "a polynucleotide of viral sequences that is replicated in host cells during the virus multiplication cycle" (Huisman *et al.*, 1992, "Genetic engineering with plant viruses", T.M.A. Wilson and



J.W.Davies eds., 1992, CRC Press, Inc.). In this invention we use the term "amplification-based expression system" to designate either a full-length viral genome or any fragment of viral RNA or DNA that (i) contains and is able to express foreign sequences, non-native for the wild-type parental virus (ii) replicates either by itself or as a result of complementation by a helper virus or by a product of the transgenic plant host. The terms "amplification-based expression system" and "recombinant viral vector" are closely similar. These systems represent a recombinant nucleic acid containing additional sequences, homologous (native) or foreign, heterologous (non-native) with respect to the viral genome. The term "non-native" means that this nucleic acid sequence does not occur naturally in the wild-type genome of the virus and originates from another virus or represents an artificial synthetic nucleotide sequence. Such an amplification-based system derived from viral elements is capable of replicating and, in many cases, cell-to-cell as well as long-distance movement either in a normal or/and in a genetically modified transgenic host plant. In the latter case the transgenic plant should complement the viral components of a vector which may be deficient in a certain function, i.e. the product(s) of a transgene essential for vector replication and/or expression of its genes or long-distance transport should be provided by the transgenic plant. Further examples of functions which may be provided by the plant are the following: amplification of the vector, virus assembly, ability to suppress the silencing of viral infection development in plant cells, ability to redirect the metabolism in plant cells, and cell-to-cell or long-distance movement of said vector.

Plant virus amplification-based vectors based on a monopartite (e.g. tobacco mosaic virus, TMV) or a multipartite (e.g. members of *Bromoviridae* family) genome have been shown to express foreign genes in host plants (for review, see "Genetic engineering with plant viruses", T.M.A.Wilson and J.W.Davies eds., 1992, CRC Press, Inc.).

The majority (about 80%) of known plant viruses contains plus-sense single-stranded RNA (ssRNA) genomes that are infectious when being isolated from the virions in a form of free RNA. This means that at the first step of the virus replication cycle, genomic RNA must be translated in order to produce the virus-specific RNA-dependent RNA polymerase (replicase) that is absent from uninfected plant cells and, therefore, is essential for viral RNA replication (for review, see Y. Okada, 1999, *Philosoph. Transact. of Royal Soc., B*, 354, 569-582). It should be mentioned that plus-sense ssRNA viruses differ in translation strategies used for genome expression: the genomes of so called picorna-like viruses represent a single continuous open reading frame (ORF) translated by the ribosome into a large polyprotein which is then proteolytically processed into functionally active virus-encoded proteins. The virus-specific proteinase(s) are involved in polyprotein processing. A second peculiar feature of picorna-like

viruses is that their genomic RNA contains, instead of a cap structure, a small viral protein covalently linked to the 5'-end of the genome.

In this invention we most preferably focus on viruses of the so-called Sindbis-like superfamily that comprises many plant viruses, in particular, more than a dozen of viruses belonging to the genus *Tobamovirus* (for review, see A.Gibbs, 1999, Philosoph. Transact. of Royal Soc., B, 354, 593-602). The technology ensures cap-independent and viral promoter-independent expression of foreign genes.

The genome of tobamoviruses (TMV U1 is the type member) contains four large ORFs. The two components of the replicase (the 130-kDa and its readthrough 183-kDa proteins) are encoded by the 5'-proximal region of the genomic RNA and are translated directly from genomic RNA. The 3'-terminal 15 nucleotides of the 180-kDa protein gene of TMV U1 overlap with the ORF coding for the 30-kDa protein responsible for cell-to-cell movement of TMV infection (movement protein, MP). In TMV U1 this gene terminates two nucleotides before the initiation codon of the last gene which encodes the 17-kDa coat protein (CP) located upstream of the 3'-proximal nontranslated region (3'-NTR) consisting of 204 nucleotides (in TMV U1).

Translation of RNA of tobamoviruses occurs by a ribosome scanning mechanism common for the majority of eukaryotic mRNAs (for reviews, see Kozak, 1989, J. Mol. Biol. 108, 229-241; Pain, 1996 ; Merrick and Hershey, 1996, In "Translational control", eds. Hershey, Matthews and Sonenberg, Cold Spring Harbour Press, pp. 31-69; Sachs and Varani, 2000, Nature Structural Biology 7, 356-360). In accordance with this mechanism, structurally polycistronic tobamovirus RNA is functionally monocistronic, i.e., only the 5'-proximal ORF encoding the replicative proteins (130-kDa protein and its readthrough product) can be translated from full-length genomic RNA (reviewed by Palukaitis and Zaitlin, 1986, In "The Plant Viruses", van Regermortel and Fraenkel-Conrat eds., vol.2, pp.105-131, Plenum Press, NY). It should be emphasized that the 68-nucleotide 5'-terminal nontranslated leader sequence of TMV U1 termed omega ( $\Omega$ ) has been shown to play the role of an efficient translational enhancer stimulating the translation of the 5'-proximal ORF.

The 5'-distal MP and CP genes are translationally silent in full-length TMV U1 RNA, however, they are translated from separate mRNAs referred to as subgenomic RNAs (sgRNA). Apparently, the tobamovirus sgRNAs are transcribed from negative-sense genomic RNA and share a common 3'-terminus. The expression of TMV genes that are translated from sgRNAs is regulated independently, both quantitatively and temporarily: the MP is produced transiently during early steps of infection and accumulates to relatively low levels (about 1% of total plant

protein), whereas the CP constitutes up to 70% of total plant protein synthesis and the CP can accumulate up to 10% of total cellular protein (Fraser, 1987, In "Biochemistry of virus-infected plants", pp.1-7, Research Studies Press Ltd., Letchworth, England).

It is clear that production of each sgRNA is controlled by different *cis*-acting sequences termed "subgenomic mRNA promoter" (sgPR). Generally, this term indicates the region of the viral genome (presumably in a minus-sense RNA copy) that can be recognized by the replicase complex to initiate transcription from the internally located sgPR sequence to produce sgRNA. However, for convenience, by the term "subgenomic promoter" we conventionally mean a nucleotide sequence in plus-sense viral RNA that is usually located upstream of the coding sequence and the start point of sgRNA and which is functionally involved in the initiation of the sgRNA synthesis. However, it should be taken into consideration that some viral sgPRs are located not only upstream of the controlled viral gene, but can even overlap with this gene (Balmori *et al.*, 1993, *Biochimie (Paris)* 75, 517-521). Each sgPR occupies a different position in the TMV genome. None of the sgPRs of TMV has been precisely mapped, but the 250 nucleotides upstream of the CP gene have been shown to promote synthesis of the CP sgRNA (Dawson *et al.*, 1989, *Virology* 172, 285-292).

Lehto *et al.* (1990, *Virology* 174, 145-157) inserted in the TMV genome (in front of the MP gene) sequences (253 and 49 nucleotides) preceding the CP gene in order to estimate the size of the CP sgPR. The insertion did not remove the native MP sgPR, but separated it from the MP ORF. The mutant (called KK6) with an inserted 253nt promoter region replicated stably and moved systemically over the infected plant. It is not unexpected that in the KK6 mutant the insertion changed the length of the MP sgRNA leader (Lehto *et al.*, 1990, *Virology* 174, 145-157) (see Fig. 9). The KK6 MP sgRNA leader was 24 nucleotides compared to 9 b.p. for the CP sgRNA.

By contrast, the mutant with an inserted 49-nt fragment of the promoter region replicated only transiently before being overtaken by a progeny of wild-type virus with the insert deleted. In addition, it has been shown (Meshi *et al.*, 1987, *EMBO J.*, 6, 2557-2563) that production of the CP sgRNA was reduced when the 96-nt region derived from CP sgPR was used. It is concluded that the 49-96nt sequences upstream of the CP gene did not contain the entire sgPR of the TMV U1 CP gene, whereas the 250-nt sequence included complete sgPR.

There is little information about the structure and mapping of sgPR controlling the expression of the TMV MP gene. Because the putative MP sgPR sequence overlaps with the 183-kDa replicase protein, the mutational analysis of the MP sgPR was complicated. Preliminary

results of W. Dawson and co-workers reported recently delineated the boundaries of the minimal and full MP sgPR of TMV U1 (Grdzlishvili *et al.*, 2000, *Virology* 276, in press). Computer folding of the region upstream of the MP gene reveals two stem-loop structures, located 5'-proximally to the 75-nt region preceding AUG codon of the MP gene.

It is assumed that in contrast to genomic RNA and the CP sgRNA, the sgRNA of the MP gene (so called I<sub>2</sub> sgRNA) is uncapped (for review see: Okada, 1999, *Philosoph. Transact. Of Royal Soc., B*, 354, 569-582). The present invention provides the results confirming the absence of the cap-structure in I<sub>2</sub> sgRNAs of both TMV U1 and crTMV (Fig. 7).

It has been shown by W. Dawson with colleagues that an important factor affecting the expression of a foreign gene from the vector virus is the position of the foreign gene relative to the 3'-terminus of viral genome: the efficiency of expression increased dramatically when the gene was placed closer to the 3'-terminus (Culver *et al.*, 1993, *Proc. Natl. Acad. Sci. USA* 90, 2055-2059). The highest expressed gene is that of the CP which is adjacent to the 3'-NTR that consists (in TMV U1 RNA) of three pseudoknots followed by a tRNA-like structure. It was suggested (Shivprasad *et al.*, 1999, *Virology* 355, 312-323) that the proximity of the gene to the pseudoknots rather than to the 3'-terminus was the main factor increasing expression of the foreign gene. Many important aspects of the TMV sg PRs structure were clarified due to the efforts of W. Dawson's group, however, the general conclusion of these authors was that "we are still in the empirical stage of vector building" (Shivprasad *et al.*, 1999, *Virology* 355, 312-323).

The above shows that the synthesis of sgRNAs is essential for expression of the 5'-distal genes of TMV genome, since these genes are translationally silent in full-length RNA. The mechanism of gene autonomization by subgenomization can be regarded as a strategy used by TMV in order to overcome the inability of eukaryotic ribosomes to initiate translation of the 5'-distal genes from polycistronic mRNA. According to the traditional ribosome scanning model (Kozak, 1999, *Gene* 234, 187-208), the internal genes of a polycistronic eukaryotic mRNA are not accessible to ribosomes.

Recently, we have isolated a crucifer infecting tobamovirus (crTMV) from *Oleracia officinalis* L. plants. A peculiar feature of crTMV was its ability to infect systemically members of *Brassicaceae* family. In addition, this virus was able to systemically infect plants of the *Solanaceae* family and other plants susceptible to TMV U1. The genome of crTMV (6312 nucleotides) was sequenced (Dorokhov *et al.*, 1994, *FEBS Letters* 350, 5-8) and was shown to

contain four traditional ORFs encoding proteins of 122-kDa (ORF1), 178-kDa (ORF2), the readthrough product of 122-kDa protein, a 30-kDa MP (ORF3), and a 17-kDa CP (ORF4). A unique structural feature of crTMV RNA was that, unlike other tobamoviruses, the coding regions of the MP and CP genes of crTMV are overlapped by 75 nucleotides, i.e. the 5'-proximal part of the CP coding region also encodes the C-terminal part of the MP.

In order to provide a clear and consistent understanding of the specification and the claims, including the scope given herein to such terms, the following definitions are provided:

**Adjacent:** A position in a nucleotide sequence immediately 5' or 3' to a defined sequence.

**Amplification vector:** A type of gene vector that, upon introduction into a host cell, is capable of replicating therein.

**Anti-Sense Mechanism:** A type of gene regulation based on controlling the rate of translation of mRNA to protein due to the presence in a cell of an RNA molecule complementary to at least a portion of the mRNA being translated.

**Chimeric Sequence or Gene:** A nucleotide sequence derived from at least two heterologous parts. The sequence may comprise DNA or RNA.

**Coding Sequence:** A deoxyribonucleotide sequence which, when transcribed and translated, results in the formation of a cellular polypeptide or a ribonucleotide sequence which, when translated, results in the formation of a cellular polypeptide.

**Compatible:** The capability of operating with other components of a system. A vector or plant viral nucleic acid which is compatible with a host is one which is capable of replicating in that host. A coat protein which is compatible with a viral nucleotide sequence is one capable of encapsidating that viral sequence.

**Gene:** A discrete nucleic acid sequence responsible for a discrete cellular product.

**Gene to be expressed:** A gene of technological interest to be expressed.

**Host:** A cell, tissue or organism capable of replicating a vector or plant viral nucleic acid and which is capable of being infected by a virus containing the viral vector or plant viral nucleic acid. This term is intended to include procaryotic and eukaryotic cells, organs, tissues or organisms, where appropriate.

**Host Plant Genome:** This term mean preferably the nuclear genome of a host plant cell, but may also include mitochondrial or chloroplast DNA.

**Infection:** The ability of a virus or amplification-based vector to transfer its nucleic acid to a host or introduce nucleic acid into a host, wherein the viral nucleic acid or a vector is replicated, viral

proteins are synthesized, and new viral particles assembled. In this context, the terms "transmissible" and "infective" are used interchangeably herein.

Internal Ribosome Entry Site (IRES) element, or IRES: a nucleotide sequence of viral, cellular or synthetic origin, which at the stage of translation is responsible for internal initiation.

IRES element necessary for translation of a gene downstream thereof: IRES element which is effective for translation of said gene in the sense that without such IRES element no technologically significant translation of this gene will occur.

Non-viral gene: A gene not functional for the life cycle of a virus.

Phenotypic Trait: An observable property resulting from the expression of a gene.

Plant Cell: The structural and physiological unit of plants, consisting of a protoplast and the cell wall.

Plant Organ: A distinct and visibly differentiated part of a plant, such as root, stem, leaf or embryo.

Plant Tissue: Any tissue of a plant in planta or in culture. This term is intended to include a whole plant, plant cell, plant organ, protoplast, cell culture, or any group of plant cells organized into a structural and functional unit.

Production Cell: A cell of a tissue or organism capable of replicating a vector or a viral vector, but which is not necessarily a host to the virus. This term is intended to include prokaryotic and eukaryotic cells, organs, tissues or organisms, such as bacteria, yeast, fungus and plant tissue.

Promoter: The 5'-non-coding sequence upstream to and operationally connected to a coding sequence which is involved in the initiation of transcription of the coding sequence.

Protoplast: An isolated plant cell without cell walls, having the potency of regeneration into cell culture or a whole plant.

Recombinant Plant Viral Nucleic Acid: Plant viral nucleic acid which has been modified to contain nonnative nucleic acid sequences.

Recombinant Plant Virus: A plant virus containing the recombinant plant viral nucleic acid.

Reporter Gene: A gene the gene product of which can be easily detected.

Subgenomic Promoter (sgPR): A promoter of a subgenomic mRNA of a vector or a viral nucleic acid.

Substantial Sequence Homology: Denotes nucleotide sequences that are homologous so as to be substantially functionally equivalent to one another. Nucleotide differences between such sequences having substantial sequence homology will be de minimus in affecting function of the gene products or an RNA coded for by such sequence.

Transcription: Production of an RNA molecule by RNA polymerase as a complementary copy of

a DNA sequence.

Translation: Production of a polypeptide by a ribosome (frequently by means of *scanning* a messenger RNA).

Vector: A nucleic acid, which is capable of genetically modifying a host cell. The vector may be single-stranded (ss) (+), ss (-) or double-stranded (ds).

Virus: An infectious agent composed of a nucleic acid encapsidated in a protein. A virus may be a mono-, di-, tri- or multi-partite virus.

### Advantages of the Invention

This invention provides a novel strategy for constructing amplification-based vectors for foreign (heterologous, non-native) gene expression such that translation of these genes can occur through an IRES-mediated internal ribosome entry mechanism from a polycistronic RNA and/or through IRES-mediated cap-independent internal ribosome entry mechanism from bi- and multicistronic sgRNA produced from the vector in the infected cell. In either event, the IRES element is necessary for translation of a gene. One of the advantages of this strategy is that it does not require any specific manipulation in terms of sgPRs: the only sequences that should be inserted into the vector are the IRES-sequence(s) (native or/and non-native) upstream of gene(s) to be translated. As a result, translation of downstream gene(s) is promoted by the inserted IRES sequences, i.e. is cap-independent. The sequence segment harboring an IRES element preferably does not function as subgenomic promoter to a technically significant degree. This means that this sequence segment either does not cause any detectable production of corresponding subgenomic RNA or that for the translation of any such subgenomic RNA, if formed by any residual subgenomic promoter activity of said sequence segment, this IRES element is still necessary for the translation of a downstream gene. Consequently, in a special case, primary recombinant RNA produced by the vector comprises: one or more structural genes preferably of viral origin, said IRES sequence, the (foreign) gene of interest located downstream of the IRES and the 3'-NTR. It is important that this strategy allows a simultaneous expression of more than one foreign gene by insertion of a tandem of two (or more) foreign genes, each being controlled by a separate IRES sequence. The present invention is preferably directed to nucleic acids and recombinant viruses which are characterised by cap- independent expression of the viral genome or of its subgenomic RNAs or of non-native (foreign) nucleic acid sequences and which are capable of expressing systemically in a host

plant such foreign sequences via additional plant-specific IRES element(s).

In a first preferred embodiment, a plant viral nucleic acid is provided in which the native coat protein coding sequence and native CP subgenomic promoter have been deleted from a viral nucleic acid, and a non-native plant viral coat protein coding sequence with upstream located plant virus IRES element has been inserted that allows for cap-independent expression in a host plant, whereas packaging of the recombinant plant viral nucleic acid and subsequent systemic infection of the host by the recombinant plant viral nucleic acid are maintained.

The recombinant plant viral nucleic acid may contain one or more additional native or non-native IRES elements that function as translation elements and which have no transcriptional activity, i.e. are effectively unable to function as a subgenomic promoter. Each native or non-native IRES element is capable of providing cap-independent expression of adjacent genes or nucleic acid sequences in the host plant.

In a second preferred embodiment, an amplification and expression vector is provided in which native or non-native plant virus IRES element(s) located upstream of foreign nucleic acid sequences are inserted downstream of a native coat protein gene. The inserted plant virus IRES element may direct cap-independent expression of adjacent genes in a host plant. Non-native nucleic acid sequences may be inserted adjacent to the IRES element such that said sequences are expressed in the host plant under translational control of the IRES element to synthesize the desired product.

In a third preferred embodiment, a recombinant vector nucleic acid is provided as in the second embodiment except that the native or non-native plant viral IRES element(s) with downstream located foreign nucleic acid sequences are inserted upstream of native coat protein subgenomic promoter and coat protein gene.

In a fourth preferred embodiment, a recombinant vector nucleic acid is provided in which native or non-native plant viral IRES element(s) is (are) used at the 5' end of the viral genome or in the viral subgenomic RNAs so as to render translation of a downstream gene(s) cap-independent.

In a fifth preferred embodiment, inhibition of cap-dependent translation is being utilised to increase the level of cap-independent translation from said vectors.

The viral-based amplification vectors are encapsidated by the coat proteins encoded by the recombinant plant viral nucleic acid to produce a recombinant plant virus. The recombinant plant viral nucleic acid is capable of replication in the host, systemic spreading in the host, and



cap-independent expression of foreign gene(s) or cap-independent expression of the whole viral genome or of subgenomic RNAs in the host to produce a desired product. Such products include therapeutic and other useful polypeptides or proteins such as, but not limited to, enzymes, complex biomolecules, or polypeptides or traits or products resulting from anti-sense RNA production.

Specific examples of proteins to be produced are antibodies, antigens, receptor antagonists, neuropeptides, enzymes, blood factors, Factor VIII, Factor IX, insulin, pro-insulin, somatotropin, serum albumin, tissue-type plasminogen activator, tissue-type plasminogen activator, haematopoietic factors such as granulocyte-macrophage colony stimulating factor, macrophage colony stimulating factor, granulocyte colony stimulating factor, interleukin 3, interleukin 11, thrombopoietin, erythropoietin, etc.

Examples for desirable input traits are resistance to herbicides, resistance to insects, resistance to fungi, resistance to viruses, resistance to bacteria, resistance to abiotic stresses, and improved energy and material utilization.

Examples for desirable output traits are modified carbohydrates, modified polysaccharides, modified lipids, modified amino acid content and amount, modified secondary metabolites, and pharmaceutical proteins, including enzymes, antibodies, antigens and the like.

Examples for trait regulation components are gene switches, control of gene expression, control of hybrid seed production, and control of apomixis.

The present invention is also directed to methods for creation of artificial, non-natural IRES elements (as opposed to IRESs isolated from living organisms) providing cap-independent and promoter independent expression of a gene of interest in plant cells (and perhaps additionally in yeast or animal cells). Artificial IRES elements may be created on the basis of the content of certain bases, notably the content of adenine and guanine bases (cf. example 14). Examples for living organisms from which IRESs may be isolated are animal viruses and plant viruses. Examples for animal viruses are hepatitis C virus, infectious bronchitis virus, picornaviruses such as poliovirus and encephalomyocarditis virus, and retroviruses such as moloney murine leukemia virus, and harvey murine sarcoma virus. Examples for plant viruses are potato virus X, potyviruses such as potato virus Y and turnip mosaic virus, tobamoviruses such as crucifer-infecting tobamovirus, and comoviruses such as cowpea mosaic virus. Alternatively, natural IRESs may be isolated from plant or animal cellular messenger RNAs like those derived from antennapedia homeotic gene, human fibroblast growth factor 2, translation

initiation factor eIF-4G, or *N. tabacum* heat shock factor 1 (see example 14). Artificial IRESes or IRESes based on IRES elements from plants and animals do not show sub-genomic promoter activity to any significant extent. Such IRESes may be used instead of the plant virus IRES elements in the embodiments described above.

In a sixth preferred embodiment, artificial, non-natural IRES elements are created on the basis of complementarity to 18S rRNA of eukaryotic cells, including yeast, animal and plant cells.

In a seventh preferred embodiment, artificial, non-natural IRES elements are created on the basis of repeated short stretches of adenosin/guanosin bases.

In an eighth preferred embodiment of this invention, a method of engineering and using viral-based amplification vectors is presented, wherein viral genome expression in plant cells occurs under the control of a plant-specific artificial transcription promoter.

In a further specific embodiment, an IRES element is used in the vector and method of the invention, which IRES element is or comprises segment(s) of a natural IRES of plant origin.

In a ninth preferred embodiment of the present invention, a method of constructing and using viral-based amplification vectors is presented, which vectors allow for expression from replicons being formed in plant cells as a result of primary nuclear transcript processing.

In a tenth preferred embodiment of this invention, a procedure is described for using circular single-stranded viral-based amplification vectors for cap-independent expression of foreign genes in plants.

In an eleventh preferred embodiment of the present invention, methods are presented that allow for expression of a gene of interest in cells under conditions favoring cap-independent translation. In one example, cells infected with an amplification vector are treated with a compound inhibiting cap-dependent translation. In another example, the vector itself contains a gene, the product of which has an inhibiting effect on cap-dependent translation in the host or an anti-sense sequence having said function.

In a twelfth preferred embodiment of this invention, a method is described that allows, by using *in vivo* genetic selection, to identify an IRES sequence that provides cap-independent expression of gene of interest or a reporter gene in an expression vector.

In a 13<sup>th</sup> embodiment, the vector of the invention is assembled from sequences derived from different viruses in order to avoid repeats of sequences in the vector and to increase the stability of the vector. This embodiment is exemplified in Example 13.

In a 14<sup>th</sup> embodiment, a gene expression system is provided comprising a vector or pro-vector according to the invention and a natural or genetically engineered plant that supports

amplification and expression of said vector. Said system preferably supports expression of a pharmaceutical protein like antibodies, antigens, receptor antagonists, neuropeptides, enzymes, blood factors, Factor VIII, Factor IX, insulin, pro-insulin, somatotropin, serum albumin, tissue-type plasminogen activator, tissue-type plasminogen activator, haematopoietic factors such as granulocyte-macrophage colony stimulating factor, macrophage colony stimulating factor, granulocyte colony stimulating factor, interleukin 3, interleukin 11, thrombopoietin, erythropoietin.

More preferably, said system provides expression of two or more genes in the same plant cell or the same plant. Alternatively, said gene expression system may further comprise an *Agrobacterium* intermediary host system that supports delivery of one or more of the vectors or pro-vectors according to the invention. Said *Agrobacterium* intermediary host may further support transfer and transient or stable expression of other traits necessary or desirable for expression of a gene to be expressed.

## BRIEF DESCRIPTION OF THE FIGURES

Fig. 1 depicts vectors T7/crTMV and SP6/crTMV.

Fig. 2 depicts vectors T7/crTMV/IRES<sub>MP,75</sub><sup>CR</sup>-GUS, T7/crTMV/IRES<sub>MP,75</sub><sup>UI</sup>-GUS, T7/crTMV/IRES<sub>MP,228</sub><sup>CR</sup>-GUS, T7/crTMV/IRES<sub>CP,148</sub><sup>CR</sup>-GUS, T7/crTMV/SPACER<sub>CP,148</sub><sup>UI</sup>-GUS and T7/crTMV/PL-GUS.

Fig. 3. Mapping of the 5'end of the crTMV I<sub>2</sub> sgRNA by primer extension (A) and putative secondary structure of I<sub>2</sub> sgRNA 5'NTR (b).

Fig. 4. crTMV I<sub>2</sub> sgRNA 5'NTR contains a translation-inhibiting hairpin structure. (A) depicts artificial transcripts used for *in vitro* translation in wheat germ extracts (WGE); (B) shows translation products synthesized in WGE.

Fig. 5. Tobamoviruses contain a putative translation-inhibiting hairpin structure upstream of the MP gene.

Fig. 6. Method of the specific detection of capped mRNAs. A, B: RNA-tag with known sequence is ligated specifically to the cap of tested RNA. C: Reverse transcription with 3'-specific primer and synthesis of first strand of cDNA. Tag sequence is included to the sequence of cDNA. D: PCR with tag-specific and 3'-specific primers. The appearance of the respective PCR band indicates the presence of cap-structure in the tested RNA. E: PCR with 5'-specific and 3'-specific primers. The appearance of PCR band serves as a control for the PCR reaction and indicates a presence of the specific tested RNA in the reaction. F: Relative comparison of the lengths of obtained PCR bands.

Fig. 7a and 7b. Detection of the presence of a cap-structure at the 5'-terminus of viral RNAs in a 2% agarose gel. Arrows indicate the respective PCR bands.

Fig. 8. depicts KK6-based TMV vectors.

Fig. 9. Nucleotide sequence of 5'NTR of KK6 and KK6-IRES<sub>MP,75</sub><sup>CR</sup> I<sub>2</sub> sgRNA.

Fig. 10. Time-course of CP and MP accumulation in leaves inoculated with KK6-IRES<sub>MP,75</sub><sup>CR</sup> (K86), KK6 and TMV UI.

Fig. 11. CP accumulation in tobacco infected with KK6, KK6-IRES<sub>MP,75</sub><sup>CR</sup>, KK6-IRES<sub>MP,125</sub><sup>CR</sup>, and KK6-H-PL and KK6-PL.

Fig. 12 depicts a crTMV IRESmp multimer structure and complementarity to 18S rRNA.

Fig. 13 depicts bicistronic transcripts containing IRES<sub>MP,75</sub><sup>CR</sup>, the tetramers of 18-nt segment of IRES<sub>CP,148</sub><sup>CR</sup>, 19-nt segment of IRES<sub>MP,75</sub><sup>CR</sup>, polylinker (PL) as intercistronic spacer and products of their translation in RRL.

Fig. 14 depicts the IRES<sub>CP,148</sub><sup>CR</sup> structure.

Fig. 15 depicts constructs used for IRES<sub>CP,148</sub><sup>CR</sup> sequence elements testing *in vitro* and *in vivo*.

Fig. 16. GUS activity testing in WGE after translation of transcripts depicted in Fig. 21.

Fig. 17. GUS activity test in tobacco protoplasts transfected with 35S promoter-based constructs analogous to those depicted in Fig. 21.

Fig. 18 depicts a scheme for cloning of two infectious TMV vectors containing IRES<sub>MP,75</sub><sup>CR</sup> in 5'NTR.

Fig. 19 depicts vector Act2/crTMV.

Fig. 20 depicts pUC-based vector Act2/crTMV/IRES<sub>MP,75</sub><sup>CR</sup>-GUS

Fig. 21 depicts circular single-stranded vector KS/Act2/crTMV/IRES<sub>MP,75</sub><sup>CR</sup>-GUS.

Fig. 22 depicts vector KS/Act2/crTMV/IRES<sub>MP,75</sub><sup>CR</sup>-GUS

Fig. 23 depicts construct 35S/CP/IRES<sub>MP,75</sub><sup>CR</sup>/GUS.

Fig. 24 depicts construct 35S/GUS/IRES<sub>MP,75</sub><sup>CR</sup>/CP.

Fig. 25 depicts construct 35S/CP-VPg/IRES<sub>MP,75</sub><sup>CR</sup>/GUS.

Fig. 26 shows a construct for *in vivo* genetic selection to identify a viral subgenomic promoter or an IRES sequence that provides cap-independent expression of a gene of interest in an expression vector.

Fig. 27 shows the restriction map of TMV-U1 cDNA clone.

Fig. 28 depicts a scheme of cloning of two infectious TMV-U1 vectors containing IRES<sub>MP,75</sub><sup>CR</sup>-GUS and IRES<sub>MP,75</sub><sup>CR</sup>-eGFP insertions.

Fig. 29 depicts vectors SP6/TMV-U1/IRES<sub>MP,75</sub><sup>CR</sup>-GUS, SP6/TMV-U1/IRES<sub>MP,75</sub><sup>UI</sup>-eGFP.

Fig. 30 shows a scheme of cloning of SP6/TMV-U1/GUS vectors containing IRES of plant origin (NtHSF) and artificial IRES ((GAAA)x16).

Fig. 31 shows the results of GUS expression from viral vectors SP6/TMV-U1/IRES<sub>MP,75</sub><sup>CR</sup>-GUS, SP6/TMV-U1/NtHSF-GUS, SP6/TMV-U1/(GAAA)X16-GUS

## DETAILED DESCRIPTION OF THE INVENTION

A primary objective of this invention is to provide a novel strategy for the construction of amplification-based vectors for foreign (heterologous, non-native) gene expression such that translation of these genes will occur by virtue of IRES-mediated cap-independent internal ribosome entry mechanism from polycistronic genomic viral RNAs and/or from bi- and multicistronic sgRNAs produced by an amplification vector, preferably, a viral vector in a plant cell.

Construction of recombinant plant viral RNAs and creation of amplification-based vectors for the introduction and expression of foreign genes in plants has been demonstrated by numerous authors using the genomes of viruses belonging to different taxonomic groups (for review, see "Genetic Engineering With Plant Viruses", 1992, eds. Wilson and Davies, CRC Press, Inc.). Tobamoviruses are considered to be convenient subjects for the construction of viral vectors. Donson *et al.* (U.S. Patents Nos. 5,316,931; 5,589,367 and 5,866,785) created TMV-based vectors capable of expressing different foreign genes in a host plant. Thus, neomycin phosphotransferase,  $\alpha$ -trichosantin and several other foreign genes were inserted adjacent to the subgenomic promoter (sgPR) of TMV CP. Donson *et al.*, (1993, PCT WO 93/03161) developed on the basis of a tobamovirus "a recombinant plant viral nucleic acid comprising a native plant viral subgenomic promoter, at least one non-native plant viral subgenomic promoter and a plant viral coat protein coding sequence, wherein said non-native plant viral subgenomic promoter is capable of initiating transcription of an adjacent nucleic acid sequence in a host plant and is incapable of recombination with the recombinant plant viral nucleic acid subgenomic promoters and said recombinant plant viral nucleic acid is capable of systemic infection in a host plant".

Contrary to the technology of Donson *et al.*, the present invention is not concerned with sgPRs in order to construct a viral replicon-based plant expression system. Instead of sgPRs, our technology manipulates with IRES-sequences of different origin (native or non-native for the virus), the sequences of which effectively lack sgPR activity, i.e. are effectively unable to promote sgRNA production. Therefore, these IRES sequences should not be regarded as sgPRs even in the case they represent a nonfunctional segment of a sgPR.

It is generally believed that uncapped transcripts of full-length viral RNA obtained after *in vitro* transcription of cDNA clones are generally non-infectious for intact plants and isolated protoplasts. Therefore, capping of a virus expression vector RNA transcript is generally

considered as a prerequisite for *in vitro* transcript infectivity. Capped RNA transcripts are commonly used for introducing a viral vector RNA into a plant. It is important to note that in some cases viral RNA may be encapsidated by the coat protein using a simple procedure of *in vitro* assembly. Thus, TMV virions as well as pseudovirions containing vector RNA can be readily produced from CP and *in vitro* transcripts or purified authentic viral RNA. About fifteen years ago, it has been shown by Meshi *et al.* (1986, Proc. Natl. Acad. Sci. USA 85, 5043-5047) that (1) the uncapped transcripts of full-length TMV RNA produced *in vitro* are infectious in the absence of a cap analogue, although their specific infectivity is very low.

In the present invention, uncapped expression vector RNA reassembled with TMV CP can be used for plant inoculations in order to overcome its low infectivity. At least one of the additional approaches described in this invention opens the technical possibilities for plant infection with a cap-independent plant viral vector. This is the method of insertion of a full-length single-stranded (ss) DNA copy of a viral vector under control of an appropriate DNA promoter. After inoculation of a host plant with the recombinant viral DNA, the infectious full-length RNA of a plant viral vector, which will be able to replicate and spread over the plant, will be produced. In other words, these procedures, taken together with the fact of cap-independent expression of foreign gene(s) promoted by IRES sequences, make both processes, namely host plant inoculation and foreign gene expression, entirely cap-independent.

An important preferred object of the present invention is the creation of a series of crTMV genome-based viral vectors with the "IRES-foreign gene" block inserted between the CP gene and 3'-NTR. Various IRES and control sequences were used (see Fig. 2) in combination with two different reporter genes (GUS and GFP). A unique feature of this invention is that the foreign genes that were located outside of the viral sgPR sequences were expressed in the infected plant cap-independently from the 3'-proximal position of genomic and sgRNAs produced by the vector. In particular, the IRES<sub>MP,75</sub><sup>CR</sup> sequence representing the 3'-terminal part of the 5'-nontranslated leader sequence of crTMV sgRNA I<sub>2</sub> was efficient in mediating cap-independent expression of the 3'-proximal foreign gene in plants infected with a viral vector. It should be emphasized that said crTMV-based viral vectors produce three types of viral plus-sense ssRNAs in infected plants, including: i) full-length genomic RNA, ii) tricistronic I<sub>2</sub> sgRNA (our data show that the latter sgRNA is uncapped, contrary to full-length RNA), and iii) bicistronic sgRNA containing the first CP gene and the second foreign gene. Therefore, all these RNAs are 3'-coterminal and cap-independent translation of their 3'-proximal gene from either

capped (full-length and bicistronic) or uncapped (tricistronic) RNAs is promoted by the preceding IRES sequence.

An important characteristic of virus-based vectors is their stability. However, the TMV-based vectors with foreign genes usually do not move efficiently through phloem in plants that can be systemically infected with wild-type virus. This may be due to increased length of the recombinant viral RNA and/or to the presence of the repeated sequences, which could lead to recombinations and deletions resulting in reversions to wild-type virus. The conversion of the progeny population to wild-type virus occurs in systemically infected leaves. A possibility to suppress such recombinations is the use of sequence elements from different origins, e.g. viral origins as exemplified in example 13.

An important characteristic for a virus-based vector is the level of foreign protein gene expression and the level of protein accumulation. The vector is able to produce readily visible bands corresponding to GUS stained in SDS-PAGE.

The technologies suitable for construction of amplification-based vectors capable of expressing foreign sequences in host plants have been developed on the basis of different viral genomes (e.g., see G. Della-Cioppa *et al.*, 1999, PCT WO 99/36516). The central feature of those inventions was that the recombinant plant viral nucleic acid "contains one or more non-native subgenomic promoters which are capable of transcribing or expressing adjacent nucleic acid sequences in the host plant. The recombinant plant viral nucleic acids may be further modified to delete all or part of the native coat protein coding sequence and to contain a non-native coat protein coding sequence under control of the native or one of the non-native plant viral subgenomic promoters, or put the native coat protein coding sequence under the control of a non-native plant viral subgenomic promoter". In other words, the most important element(s) of that invention is/are the native and non-native sgPR sequences used for artificial sgRNA production by the viral vector. An important feature that distinguishes the present invention from others is that according to WO 99/36516, the foreign gene must be inevitably located directly downstream of the sgPR sequence, i.e. should be located at the 5'-proximal position of the chimeric sgRNA produced by the viral vector in the host plant. By contrast, our invention proposes that the foreign gene is separated from a sgPR (if present) at least by one (or more) viral gene(s) such that said foreign gene is located 3'-proximally or internally within the functionally active chimeric sgRNA produced by the vector. Thus, foreign gene expression is promoted by the IRES sequence, native or non-native, of the wild-type virus.

The next preferred object of this invention is the construction of a novel type of non-



native IRES sequences, namely artificial, non-natural synthetic IRESs capable of promoting cap-independent translation of 5'-distal genes from eukaryotic polycistronic mRNAs. We show that intercistronic spacers complementary to 18S rRNA of varying length and composition are able to mediate cap-independent translation of the 3'-proximal GUS gene in bicistronic H-GFP-IRES-GUS mRNA (Fig. 13). Further, gene expression under translational control of an artificial IRES element having a high adenine nucleotide content is demonstrated using an IRES element consisting of 16 copies of the GAAA segment (example 14).

The last but not least advantage provided by the present invention is the possibility to combine repeats of two or more foreign genes each being preceded by the native or non-native IRES sequence in the amplification-based vector genome. Expression of such a cassette of an "IRES-foreign gene" will allow the simultaneous production of two or more foreign proteins by the vector.

Viruses belonging to different taxonomic groups can be used for the construction of virus-based vectors according to the principles of the present invention. This is right for both RNA- and DNA-containing viruses, examples for which are given in the following (throughout this document, each type species name is preceded by the name of the order, family and genus it belongs to. Names of orders, families and genera are in italic script, if they are approved by the ICTV. Taxa names in quotes (and not in italic script) indicate that this taxon does not have an ICTV international approved name. Species (vernacular) names are given in regular script. Viruses with no formal assignment to genus or family are indicated):

#### **DNA Viruses:**

**Circular dsDNA Viruses:** Family: *Caulimoviridae*, Genus: *Badnavirus*, type species: *commelina yellow mottle virus*, Genus: *Caulimovirus*, Type species: *cauliflower mosaic virus*, Genus "*SbCMV-like viruses*", Type species: *Soybean chlorotic mottle virus*, Genus "*CsVMV-like viruses*", Type species: *Cassava vein mosaic virus*, Genus "*RTBV-like viruses*", Type species: *Rice tungro bacilliform virus*, Genus: "*Petunia vein clearing-like viruses*", Type species: *Petunia vein clearing virus*;

**Circular ssDNA Viruses:** Family: *Geminiviridae*, Genus: *Mastrevirus* (Subgroup I Geminivirus), Type species: *maize streak virus*, Genus: *Curtovirus* (Subgroup II Geminivirus), Type species: *beet curly top virus*, Genus: *Begomovirus* (Subgroup III Geminivirus), Type species: *bean golden mosaic virus*;

#### **RNA Viruses:**

**ssRNA Viruses:** Family: Bromoviridae, Genus: Alfavirus, Type species: alfalfa mosaic virus, Genus: Ilarvirus, Type species: tobacco streak virus, Genus: Bromovirus, Type species: brome mosaic virus, Genus: Cucumovirus, Type species: cucumber mosaic virus;

Family: Closteroviridae, Genus: Closterovirus, Type species: beet yellows virus, Genus: Crinivirus, Type species: Lettuce infectious yellows virus, Family: Comoviridae, Genus: Comovirus, Type species: cowpea mosaic virus, Genus: Fabavirus, Type species: broad bean wilt virus 1, Genus: Nepovirus, Type species: tobacco ringspot virus;

Family: Potyviridae, Genus: Potyvirus, Type species: potato virus Y, Genus: Rymovirus, Type species: ryegrass mosaic virus, Genus: Bymovirus, Type species: barley yellow mosaic virus;

Family: Sequiviridae, Genus: Sequivirus, Type species: parsnip yellow fleck virus, Genus: Waikavirus, Type species: rice tungro spherical virus;

Family: Tombusviridae, Genus: Carmovirus, Type species: carnation mottle virus, Genus: Dianthovirus, Type species: carnation ringspot virus, Genus: Machlomovirus, Type species: maize chlorotic mottle virus, Genus: Necrovirus, Type species: tobacco necrosis virus, Genus: Tombusvirus, Type species: tomato bushy stunt virus, Unassigned Genera of ssRNA viruses, Genus: Capillovirus, Type species: apple stem grooving virus;

Genus: Carlavirus, Type species: carnation latent virus;

Genus: Enamovirus, Type species: pea enation mosaic virus;

Genus: Furovirus, Type species: soil-borne wheat mosaic virus, Genus: Hordeivirus, Type species: barley stripe mosaic virus, Genus: Idaeovirus, Type species: raspberry bushy dwarf virus;

Genus: Luteovirus, Type species: barley yellow dwarf virus;

Genus: Marafivirus, Type species: maize rayado fino virus;

Genus: Potexvirus, Type species: potato virus X;

Genus: Sobemovirus, Type species: Southern bean mosaic virus, Genus: Tenuivirus, Type species: rice stripe virus;

Genus: Tobamovirus, Type species: tobacco mosaic virus;

Genus: Tobravirus, Type species: tobacco rattle virus;

Genus: Trichovirus, Type species: apple chlorotic leaf spot virus;

Genus: Tymovirus, Type species: turnip yellow mosaic virus;

Genus: Umbravirus, Type species: carrot mottle virus;

Negative ssRNA Viruses: Order: Mononegavirales, Family: Rhabdoviridae, Genus: Cytorhabdovirus, Type Species: lettuce necrotic yellows virus, Genus: Nucleorhabdovirus, Type species: potato yellow dwarf virus;

Negative ssRNA Viruses: Family: Bunyaviridae, Genus: Tospovirus, Type species: tomato spotted wilt virus;

**dsRNA Viruses:** Family: Partitiviridae, Genus: Alphacryptovirus, Type species: white clover cryptic virus 1, Genus: Betacryptovirus, Type species: white clover cryptic virus 2, Family: Reoviridae, Genus: Fijivirus, Type species: Fiji disease virus, Genus: Phytoreovirus, Type species: wound tumor virus, Genus: Oryzavirus, Type species: rice ragged stunt virus;

**Unassigned Viruses:** Genome ssDNA: Species banana bunchy top virus, Species coconut foliar decay virus, Species subterranean clover stunt virus,

Genome dsDNA, Species cucumber vein yellowing virus,

Genome dsRNA, Species tobacco stunt virus,

Genome ssRNA, Species Garlic viruses A,B,C,D, Species grapevine fleck virus, Species maize white line mosaic virus, Species olive latent virus 2, Species ourmia melon virus, Species Pelargonium zonate spot virus;

**Satellites and Viroids:** Satellites: ssRNA Satellite Viruses: Subgroup 2 Satellite Viruses, Type species: tobacco necrosis satellite,

Satellite RNA, Subgroup 2 B Type mRNA Satellites, Subgroup 3 C Type linear RNA Satellites, Subgroup 4 D Type circular RNA Satellites,

Viroids, Type species: potato spindle tuber viroid.

In particular, the methods of the present invention can preferably be applied to the construction of virus replicon-based vectors using the recombinant genomes of plus-sense ssRNA viruses preferably belonging to the genus *Tobamovirus* or to the families *Bromoviridae* or *Potyviridae* as well as DNA-containing viruses. In the latter case the foreign gene should preferably be located downstream of a viral gene and its expression can be mediated by the IRES sequence from bicistronic or polycistronic mRNA transcribed by a DNA-dependent RNA polymerase from a genomic transcription promoter.

A separate preferred aspect of this invention is concerned with the application of the methods of the invention to the construction of ssDNA-based vectors. The geminivirus-based vectors expressing the foreign gene(s) under control of an IRES sequence can exemplify this aspect. The geminiviruses represent a group of plant viruses with monopartite or bipartite circular ssDNA that have twinned quasiicosahedral particles (reviewed by Hull and Davies, 1983, *Adv. Virus Res.* 28, 1-45; Mullineaux *et al.*, 1992, "Genetic engineering with plant viruses", Wilson and Davies, eds., 1992, CRC Press, Inc.). The two ssDNA components of the bipartite geminiviruses referred to as A and B encode for 4 and 2 proteins, respectively. The DNA A contains the CP gene and three genes involved in DNA replication, whereas the DNA B encodes two proteins essential for viral movement. It has been demonstrated that the genomes of bipartite geminiviruses belonging to the genus *Begomovirus*, such as tomato golden mosaic virus (TGMV) and bean golden mosaic virus (BGMV) can replicate and spread over a certain host plant despite the deletion of the CP gene (Gardiner *et al.*, 1988, *EMBO J.* 7, 899-904; Jeffrey *et al.*, 1996, *Virology* 223, 208-218; Azzam *et al.*, 1994, *Virology* 204, 289-296). It is noteworthy that some begomoviruses including BGMV exhibit phloem-limitation and are restricted to cells of the vascular system. Thus, BGMV remains phloem-limited, while TGMV is capable of invading the mesophyll tissue in systemically infected leaves (Petty and Morra, 2000, Abstracts of 19<sup>th</sup> Annual meeting of American Society for Virology, p.127).

The present invention proposes to insert the foreign gene in a bipartite geminivirus genome by two ways: (i) downstream of one of its (e.g., BGMV) genes, in particular the CP gene such that the CP ORF will be intact or 3'-truncated and the IRES sequence will be inserted upstream of the foreign gene. Therefore, the mRNA transcription will proceed from the native DNA promoter resulting in production of bicistronic chimeric mRNA comprising the first viral gene (or a part thereof), the IRES sequence and the 3'-proximal foreign gene expression of which is mediated by the IRES. Alternatively (ii), the full-length DNA copy of the the RNA genome of the viral vector can be inserted into a DNA of a CP-deficient bipartite geminivirus under control of the CP gene promoter. The RNA genome of the RNA-vector-virus will be produced as a result of DNA A transcription in the plant cell inoculated with a mixture of recombinant DNA A and unmodified DNA B. An advantage of this method is that the geminivirus-vector is needed as a vehicle used only for delivering the vector to primary-inoculated cells: all other steps will be performed by a tobamovirus vector itself including production of IRES-carrying vector RNA after geminivirus-vector DNA transcription by a cellular RNA polymerase, its replication, translation and systemic spread over the host plant and foreign

gene(s) expression. As an additional possibility for the creation of a ssDNA vector, cloning of the viral cDNA and the foreign gene into a phagemid vector and production of the ssDNA according to standard methods can be mentioned.

Taking into account that tobamovirus-derived IRES sequences are shown to be functionally active in animal cells (our previous patent application), the methods of the present invention can be used for constructing the recombinant viral RNAs and producing the viral vectors on the basis of animal viruses, e.g. the viruses belonging to the families *Togaviridae*, *Caliciviridae*, *Astroviridae*, *Picornaviridae*, *Flaviviridae* in order to produce new vectors expressing the foreign genes under control of plant virus-derived IRES sequences. Such animal virus-based vectors for plants and animals can be useful in the fields of vaccine production or for gene therapy.

It should be noted, however, that the rod-like virions of Tobamoviruses and, in particular, the flexible and long virions of filamentous Potexviruses, Carlaviruses, Potyviruses and Closteroviruses apparently provide the best models for realization of the methods of the present invention.

In another embodiment of this invention, the IRES sequence is used in such a way that the virus-based amplification vector will contain the IRES-sequence within its 5'-NTR. It is presumed that insertion of an IRES sequence does not prevent viral replication, but is able to ensure an efficient cap-independent translation of transcripts of genomic vector RNA. Therefore, said construct may comprise: (i) An IRES element within or downstream of the 5'-untranslated leader sequence that is native or non-native for said viral vector and promotes cap-independent translation of the viral 5'-proximal gene (the RdRp), and (ii) at least one native or non-native IRES sequence located downstream of one or more viral structural genes and upstream of foreign gene(s) in order to promote their cap-independent translation. According to this method, the specific infectivity of uncapped full-length vector transcripts will be increased due to efficient 5'-IRES-mediated translation of the parental RNA molecules in the primary inoculated cells.

A further preferred embodiment is a method of producing one or several protein(s) of interest in plant cells based on the introduction and cap-independent expression of a foreign gene from a mono- or polycistronic mRNA sequence mediated by the plant specific IRES sequence located upstream of said foreign gene sequence. A particular feature of this method is that the technology involves a procedure that allows to selectively switch off the cellular cap-dependent mRNA translation with the help of certain chemical compounds. However, this

procedure does not affect the cap-independent IRES-mediated translation of mRNAs artificially introduced in the plant cells, thus allowing to control and enhance cap-independent expression.

Alternatively, the means for inhibiting the translation of cellular capped mRNA can be applied to plants infected with said viral vector itself that expresses the foreign gene(s) in a cap-independent manner. Under conditions when the translation of the cellular capped mRNAs is prevented, selective expression of the foreign gene(s) from said virus vector will occur.

The vector of the invention may be an RNA or DNA vector. It may be ss(+), ss(-) or ds. It may show any of the modes of amplification known from viruses. This includes the multiplication of the vector nucleic acid and optionally the production of coat protein and optionally the production of proteins for cell-to-cell movement or long-distance movement. The genes for the required replication and/or coat and/or movement may be wholly or partially encoded in an appropriately engineered host plant. In this manner, a system is generated consisting of mutually adapted vector and host plant.

The vector may be derived from a virus by modification or it may be synthesized *de novo*. It may have only IRES elements effectively devoid of any subgenomic promoter activity. However, the vector may combine one or several subgenomic promoters with one or several IRES elements effectively devoid of subgenomic promoter function, so that the number of cistrons is greater than the number of promoters.

Considering the simplest case of one IRES element, said element may be located upstream of a (foreign) gene of interest to be expressed directly by said IRES element and optionally downstream of a (viral) gene for, say replication, to be expressed IRES-independent. Alternatively, the gene of interest may be upstream of an IRES element and expressed IRES-independent and the IRES element serves for the expression of a downstream viral gene. These simplest cases may of course be incorporated singly or multiply in a more complex vector.

The vector may contain a sequence in anti-sense orientation for suppressing a host gene. This suppression function may exist alone or in combination with the expression of a (foreign) gene of interest. A particularly preferred case involves the suppression of a gene essential for cap-dependent translation, e.g. a gene for a translation initiation factor (e.g. eIF4) associated with cap-dependent translation, so that the translation machinery of the host plant is wholly in service of vector gene translation. In this case, the vector must be wholly cap-independent. Of course, the vector may be generated within a plant cell from a pro-vector by the plant nucleic acid processing machinery, e.g. by intron splicing.

It is possible to increase the expression level of a foreign or viral gene that is translated via IRES by inhibiting the post-transcriptional gene silencing (PTGS). One of the methods is co-expression of so called anti-silencing proteins together with the protein of interest (for example, HC-Pro from tobacco etch virus or 19K protein coded by tomato bushy stunt virus, see Kasschau and Carrington, 1998, *Cell*, 95, 461-470; Voinnet *et al.*, *Proc. Natl. Acad. Sci. USA*, 96, n24, 14147-14152). Inhibitors of PTGS might be expressed either stably (transgenic plant) or transiently (viral vector, agroinoculation).

Proteins that are expressed from IRES-based vectors might be also modified using mechanisms of post-translational modifications supported by a host plant like glycosylation or proteolytic cleavage and others.

The IRES element may be of plant viral origin. Alternatively, it may be of any other viral origin as long as it satisfies the requirement of operation in a plant cell. Further, an IRES element operative in a plant cell may be a synthetic or an artificial element. Synthesis may be guided by the sequence of the 18S rRNA of the host plant, namely the segment operative for IRES binding. It should be sufficiently complementary thereto. Sufficiency of complementarity can simply be monitored by testing for IRES functionality. Complementarity in this sense comprises GC, AU and to some extent GU base pairing. Further, such IRES element may be a multimer of such a complementary sequence to increase efficiency. The multimer may consist of identical essentially complementary sequence units or of different essentially complementary sequence units. Moreover, artificial IRES elements with high translation efficiency and effectively no subgenomic promoter activity may be generated by a process of directed evolution (as described e.g. in US 6,096,548 or US 6,117,679). This may be done *in vitro* in cell culture with a population of vectors with IRES element sequences that have been randomized as known per se. The clones which express a reporter gene operably linked to the potential IRES element are selected by a method known per se. Those clones which show subgenomic promoter activity are eliminated. Further rounds of randomization and selection may follow.

The IRES element of the vector of the invention may be effectively devoid of promoter activity. This means that that the expression of a gene operably linked to an IRES element would not occur by a residual subgenomic promoter activity. This mode of action may be determined by standard molecular biology methods such as Northern blotting, primer extension analysis (Current Protocols in Molecular Biology, Ed. By F. Ausubel et al., 1999, John Wiley & Sons), 5' RACE technology (GibcoBRL, USA), and alike. It should be added that IRES elements that show detectable subgenomic promoter activity but operate essentially as translational rather

than transcriptional elements, are also subject of our invention. Such discrimination could be derived, for example, by measuring quantitatively the relative amounts of two types of mRNAs on Northern blots, namely the short mRNA due to sgPR activity and the long mRNA not due to sgPR activity. If the IRES element does not essentially operate as a residual viral subgenomic promoter, the relative amount of corresponding short mRNA should be lower than 20%, preferably lower than 10% and most preferably lower than 5% of the sum of the short and long mRNA. Thus we provide as a preferred embodiment a vector capable of amplification of a gene in a plant comprising a nucleic acid having a sequence for at least one non-viral gene to be expressed and having or coding for at least one IRES element necessary for translation of said gene in said plant with the proviso that the expression of said gene is essentially derived from translational rather than transcriptional properties of said IRES element sequence when measured by standard procedures of molecular biology.

The novel vectors of the invention open new avenues for genetic modification of plants. As a first possibility we suggest the use for determining the function of a structural gene of a plant. This is notably of interest for genomics. Therefore, a plant for which the genome has been sequenced is of particular interest. This is a small scale (plant-by-plant) application. The vector of this invention is highly effective for this application, since it allows suppression of genes of interest and/or overexpression of genes to bring out the gene function to be discovered in an intensified manner.

In a large scale application the vector may be used to generate a trait or to produce a protein in a host plant. Infection of plants with the vector may be done on a farm field previously planted with unmodified plants. This allows for the first time a genetic modification of plants on a field, whereby the farmer has greatest freedom in terms of selection of seeds and vectors from a variety of sources for producing a desired protein or trait.

Examples for plant species of interest for the application of this invention are monocotyledonous plants like wheat, maize, rice, barley, oats, millet and the like or dicotyledonous plants like rape seed, canola, sugar beet, soybean, peas, alfalfa, cotton, sunflower, potato, tomato, tobacco and the like.

In the following, the invention will be further described using specific examples. Standard molecular biological techniques were carried out according to Sambrook *et al.* (1989, Molecular Cloning: a Laboratory Manual. 2nd edn. Cold Spring Harbor Laboratory, Cold Spring Harbor,



New York). All plasmids utilized in the invention can be prepared according to the directions of the specification by a person of ordinary skill in the art without undue experimentation employing materials readily available in the art.

**EXAMPLE 1****Construction of a tobamovirus vector infecting cruciferous plants**

Virions of a known tobamovirus called crucifer tobamovirus (crTMV) which is able to infect systemically crucifer plants were isolated from *Olearacia officinalis* L. with mosaic symptoms. Results of crTMV host-range examination are presented in Table1.

**Plasmid constructions**

CrTMV cDNA was characterized by dideoxynucleotide sequencing (Dorokhov *et al.*, 1994 FEBS Letters 350, 5-8). Full length T7 RNA polymerase promoter-based infectious crTMV cDNA clones were obtained by RT-PCR from crTMV RNA using

oligonucleotides crTMV1 - Kpn 5' -  
 gcat***ggtacc***ccttaatacgactcactataGTTTTAGTTTTATTGCAACAACAACAA

(upstream), wherein the italic bold letters are a sequence of a Kpn I site, the underlined lowercase letters are nucleotide sequence of the T7 RNA polymerase promoter, the uppercase letters are from the 5'-termini of crTMV cDNA; and crTMV2 5'-  
 gcat***gcgccgc***TGGGCCCCTACCCGGGGTTAGGG (downstream), wherein the italic bold letters are sequence of NotI site, the uppercase letters are from 3'-termini of crTMV cDNA and cloning into pUC19 between KpnI and Bam HI restriction sites (Fig. 1).

Full length SP6 RNA polymerase promoter-based infectious crTMV cDNA clones were obtained by RT-PCR from crTMV RNA by using oligonucleotides crTMV1-SP6 5'-  
 gcat***ggtacc***atttaggtgacactatagaactcGTTTTAGTTTTATTGCAACAACAACAA (upstream), wherein the italic bold letters are a sequence of a Kpn I site, the underlined lowercase letters are a nucleotide sequence of the T7 RNA polymerase promoter, the uppercase letters are from the 5'-termini of crTMV cDNA; and crTMV2 5'-gcat***gcgccgc***TGGGCCCCTACCCGGGGTTAGGG (downstream), wherein the italic bold letters are a sequence of a Not I site, the uppercase letters are from 3'-termini of crTMV cDNA and cloning into pUC19 between KpnI and Bam HI restriction sites (Fig. 1).

The full-length crTMV cDNA clones were characterized by dideoxynucleotide sequencing. The ability of crTMV infectious transcripts to infect systemically *Nicotiana* and crucifer species was confirmed by infection tests on respectively *Nicotiana tabacum* var. Samsun and *Arabidopsis thaliana*.

TABLE 1. Virus detection and symptoms caused by crTMV in mechanically infected plants.

Species	Inoculated Leaves		Non-inoculated Upper Leaves	
	Symptoms <sup>*</sup>	Virus <sup>**</sup>	Symptoms	Virus
<i>Nicotiana tabacum</i> L.				
cv. Samsun	C	+	M	+
cv. Samsun NN.	L	+	s	-
<i>Nicotiana clevelandii</i> L.	L+N	+	M	+
<i>Nicotiana glutinosa</i> L.	L+N	+	s	-
<i>Nicotiana sylvestris</i> L.	L+N	+	s	+
<i>Nicotiana benthamiana</i> L.	L+N	+	M	+
<i>Nicotiana rustica</i> L.	C	+	M	+
<i>Lycopersicum esculentum</i> L.	L+N	+	s	-
<i>Solanum tuberosum</i> L.	s	-	s	-
<i>Capsicum frutescens</i> L.	L+N	+	M	+
<i>Brassica chinensis</i> L.	C	+	M	+
<i>Brassica rapa</i> L.	C	+	M	+
<i>Brassica napus</i> L.	C	+	M	+
<i>Brassica oleracea</i> L.	L	+	s	-
<i>Brassica compestris</i> L.	C	+	M	+
<i>Brassica cauliflora</i> L.	C	+	s	-
<i>Arabidopsis thaliana</i> L.	L+N	+	M	+
<i>Chenopodium amaranticolor</i> L.	L+N	+	s	+
Coste and Reyn.				
<i>Chenopodium quinoa</i> L. Willd.	L+N	+	s	-
<i>Chenopodium murale</i> L.	L+N	+	s	-
<i>Datura stramonium</i> L.	L+N	+	s	-
<i>Plantago major</i> L.	L+N	+	M	+
<i>Tetragonia expansa</i> L.	L+N	+	s	-
<i>Beta vulgaris</i> L.	L+N	+	s	-
<i>Petunia hybrida</i> L.	C	+	M	+
<i>Cucumis sativus</i> L.	L+N	+	s	-
<i>Phaseolus vulgaris</i> L.	s	-	s	-
<i>Raphanus sativus</i> L.	s	-	s	-
<i>Sinapis alba</i> L.	C	+	M	+

<sup>\*</sup>C, chlorosis; L, local lesion; M, mosaic; N, necrosis; s, symptomless.

<sup>\*\*</sup>Virus detected (+) or not (-) by ELISA.

**EXAMPLE 2****Construction of tobamoviral vectors for expression of GUS genes in *Nicotiana* and crucifer plants via viral IRESs**

Series of IRES-mediated expression vectors T7/crTMV/GUS were constructed as follows. First, Hind III and Xba I sites were inserted in the end of the CP gene of Sac II/Not I fragment of T7/crTMV vector (Fig. 1) by a polymerase chain reaction (PCR) and two pairs of specific primers. Second, IRES<sub>MP,75</sub><sup>CR</sup>-GUS, IRES<sub>MP,75</sub><sup>UI</sup>-GUS, IRES<sub>MP,228</sub><sup>CR</sup>-GUS, IRES<sub>CP,148</sub><sup>CR</sup>-GUS, IRES<sub>CP,148</sub><sup>UI</sup>-GUS, PL-GUS cDNA described in Skulachev *et al.* (1999, *Virology* 263, 139-154) were inserted into the Hind III and Xba I containing Sac II/Not I fragment of the T7/crTMV vector to obtain Sac II-IRES<sub>MP,75</sub><sup>CR</sup>-GUS-Not I, Sac II-IRES<sub>MP,75</sub><sup>UI</sup>-GUS-Not I, Sac II-IRES<sub>MP,228</sub><sup>CR</sup>-GUS-Not I, Sac II-IRES<sub>CP,148</sub><sup>CR</sup>-GUS-Not I, Sac II-IRES<sub>CP,148</sub><sup>UI</sup>-GUS-Not I, Sac II-PL-GUS-Not I cDNA, respectively. Third, Sac II-Not I cDNA fragment of T7/crTMV vector was replaced by Sac II-IRES<sub>MP,75</sub><sup>CR</sup>-GUS-Not I or Sac II-IRES<sub>MP,75</sub><sup>UI</sup>-GUS-Not I or Sac II-IRES<sub>MP,228</sub><sup>CR</sup>-GUS-Not I or Sac II-IRES<sub>CP,148</sub><sup>CR</sup>-GUS-Not I or Sac II-IRES<sub>CP,148</sub><sup>UI</sup>-GUS-Not I or Sac II-PL-GUS-Not I cDNA to obtain vector T7/crTMV/IRES<sub>MP,75</sub><sup>CR</sup>-GUS (Fig. 2), vector T7/crTMV/IRES<sub>MP,75</sub><sup>UI</sup>-GUS (Fig. 2), vector T7/crTMV/IRES<sub>MP,228</sub><sup>CR</sup>-GUS (Fig. 2), vector T7/crTMV/IRES<sub>CP,148</sub><sup>CR</sup>-GUS (Fig. 2), vector T7/crTMV/IRES<sub>CP,148</sub><sup>UI</sup>-GUS (Fig. 2) and vector T7/crTMV/PL-GUS (Fig. 2), respectively.

**EXAMPLE 3****Expression of GUS gene in transfected *Nicotiana* and crucifer plants via viral IRESs**

This example demonstrates tobamovirus IRES-mediated expression of the GUS gene in *Nicotiana benthamiana* and *Arabidopsis thaliana* plants infected crTMV-based vectors: T7/crTMV/IRES<sub>MP,75</sub><sup>CR</sup>-GUS (Fig. 2), vector T7/crTMV/IRES<sub>MP,75</sub><sup>UI</sup>-GUS (Fig. 2), vector T7/crTMV/IRES<sub>MP,228</sub><sup>CR</sup>-GUS (Fig. 2), vector T7/crTMV/IRES<sub>CP,148</sub><sup>CR</sup>-GUS (Fig. 2), vector T7/crTMV/IRES<sub>CP,148</sub><sup>UI</sup>-GUS (Fig. 2) and vector T7/crTMV/PL-GUS (Fig. 2).

***In vitro* transcription**

The plasmids T7/crTMV/IRES<sub>MP,75</sub><sup>CR</sup>-GUS (Fig. 2), vector T7/crTMV/IRES<sub>MP,75</sub><sup>UI</sup>-GUS (Fig. 2), vector T7/crTMV/IRES<sub>MP,228</sub><sup>CR</sup>-GUS (Fig. 2), vector T7/crTMV/IRES<sub>CP,148</sub><sup>CR</sup>-GUS (Fig. 2), vector T7/crTMV/IRES<sub>CP,148</sub><sup>UI</sup>-GUS (Fig. 2) and vector T7/crTMV/PL-GUS (Fig. 2) were linearized by Not I. The recombinant plasmids were transcribed *in vitro* as described by Dawson *et al.* (1986 *Proc. Natl. Acad. Sci. USA* 83, 1832-1836). Agarose gel electrophoresis of RNA transcripts confirmed

that they were intact. The RNA concentration was quantified by agarose gel electrophoresis and spectrophotometry.

#### GUS detection

Inoculated leaves were collected 10-14 days after transfection with capped full-length transcripts. IRES activity was monitored by histochemical detection of GUS expression as described earlier (Jefferson, 1987, *Plant Molecular Biology Reporter* **5**, 387-405). Samples were infiltrated using the colorimetric GUS substrate, but the method (De Block and Debrouwer, 1992, *Plant J.* **2**, 261-266) was modified to limit the diffusion of the intermediate products of the reaction: 0.115 M phosphate buffer, pH 7.0 containing 5-bromo-4-chloro-3-indolyl- $\beta$ -D-glucuronide (X-Gluc) 600  $\mu$ g/ml; 3 mM potassium ferricyanide; 10 mM EDTA. After incubation overnight at 37°C, the leaves were destained in 70% ethanol and examined by light microscopy.

#### EXAMPLE 4

IRES<sub>MP,75</sub><sup>CR</sup> does not function as MP subgenomic promoter but provides MP gene expression via cap-independent internal initiation of translation in TMV-infected plants

This example uses different approaches to confirm the possibility of IRES<sub>MP,75</sub><sup>CR</sup> used in viral vectors for cap-independent expression of a gene of interest.

CrTMV MP subgenomic RNA has a 125-nt long 5'-nontranslated region (5'NTR) and contains a translation inhibiting stem-loop secondary structure.

To determine the length and nucleotide sequence of TMV UI and crTMV MP subgenomic RNA (I<sub>2</sub> sgRNA) 5'NTR, the protocol of primer extension experiments described by Lehto *et al.* (1990, *Virology* **174**, 145-157) was changed in the following way: (i) AMV reverse transcriptase (RT); (ii) RT reaction under 45°C; (iii) the GC-rich primer; (iv) increased dNTP concentration; (v) dITP to avoid secondary structure. It has been shown (Fig. 3) that the 5'UTR sequence of crTMV I<sub>2</sub> sgRNAs consists of 125 nucleotides. This result was confirmed by direct 5'UTR RT sequencing. Fig. 3B shows that crTMV 5'NTR contains a stable hairpin-loop structure. Being placed just upstream of the MP gene of artificial transcript, it is able to inhibit MP gene translation *in vitro* (Fig. 4). This means that IRES<sub>MP,75</sub><sup>CR</sup> located between 5'HI<sub>2</sub><sup>CR</sup> and the MP gene can provide efficient cap-independent internal initiation of translation. Fig. 5 shows that

homologous to 5'HI<sub>2</sub><sup>CR</sup> putative translation inhibiting hairpin-loop structure can be revealed in the 125-nt sequence upstream of the MP gene of other tobamoviruses.

#### CrTMV and TMV UI MP subgenomic RNAs are not capped

To study the structure of the 5'-terminus of the subgenomic RNA coding for the 30K movement protein (MP) gene of crTMV, the "Jump-Start" method offered by Active Motif was used. Jump-Start<sup>TM</sup> is the method of chemical ligation of an RNA tag specifically to the 5'-end of capped mRNAs. During reverse transcription, the ribo-oligonucleotide tag of a known sequence becomes incorporated into the 3'-end of a first strand cDNA. This creates a known priming site suitable for PCR.

Initially, the 5'-terminal 2'-3'-cis-glycol groups of capped RNA were converted to reactive di-aldehydes via sodium periodate oxidation. 1-2 µl of a tested RNA (1 µg/µl) were mixed with 14 µl of pure water and 1 µl of sodium acetate buffer (pH 5.5), then 4 µl of 0.1 M sodium periodate were added and the reaction mixture was incubated for 1 hour.

Then a 3'-aminoalkyl derivatized synthetic ribo-oligonucleotide tag was chemically ligated to the di-aldehyde ends of oxidized RNA via reductive amination in the presence of sodium cyanoborohydride. 5 µl of sodium hypophosphite were added and the reaction mixture was incubated for 10 minutes. Then 23 µl of water, 1 µl of sodium acetate buffer (pH 4.5) and 2 µl of ribo-oligonucleotide tag 5'-CTAATACGACTCACTATAGGG (28,5 pmol/µl) were added to the reaction mixture and incubated for 15 minutes. Then 10 µl of sodium cyanoborohydride were added and incubated for 2 hours. Then 400 µl of 2 % lithium perchlorate in acetone were added, incubated for 15 minutes at -20°C and centrifugated for 5 minutes. The pellet was washed with acetone twice, then dissolved in 20 µl of water.

To remove an abundance of the RNA tag, CTAB precipitation in the presence of 0.3 M NaCl was used. CTAB is a strong cationic detergent that binds to nucleic acids to form an insoluble complex. Complex formation is influenced by the salt concentration: when the salt concentration is above 1 M, no complex formation occurs; when it is below 0.2 M, all nucleic acids are efficiently included in the complex; and when between 0.3 M and 0.4 M, the incorporation of small single-stranded nucleic acids into the complex is very inefficient (Belyavsky *et al.*, 1989, *Nucleic Acids Res.* 25, 2919-2932; Bertoli *et al.*, 1994, *BioTechniques* 16, 1054-1058). 10 µl of 1.2 M NaCl (to a final concentration of 0.4 M) and 3 µl of 10% CTAB (to a final concentration of 1%) were added, the reaction mixture was incubated for 15 minutes at room temperature and then centrifugated for 5 minutes. The pellet was resuspended in 10 µl of

NaCl, 20  $\mu$ l of water and 3  $\mu$ l 10% CTAB were added and the reaction mixture was incubated for 15 minutes at room temperature and then centrifugated for 5 minutes. The pellet was dissolved in 30  $\mu$ l of 1.2 M NaCl, 80  $\mu$ l of 96% ethanol was added, and the reaction mixture was incubated overnight at  $-20^{\circ}\text{C}$ . Then it was centrifugated for 5 minutes and washed with 70% ethanol. Then the pellet of tagged RNA was dissolved in 24  $\mu$ l of water.

Finally, reverse transcription with 3'-gene specific primers resulted in incorporation of the 5'-tag sequence at the 3'-terminus of first-strand cDNA. For reverse transcription, 12  $\mu$ l of tagged RNA, 1  $\mu$ l of specific 3'-end primers, 4  $\mu$ l of 5x buffer for SuperScript™ II (Gibco BRL Life Technologies) containing 250 mM Tris-HCl (pH 8.3), 375 mM KCl, 15 mM  $\text{MgCl}_2$  were mixed and heated at  $95^{\circ}\text{C}$  for 30 seconds, then cooled on ice. Then to the reaction mixture 0.5  $\mu$ l of DTT (to 1 mM final concentration), 2  $\mu$ l of 10 mM dNTP, 0.5  $\mu$ l of RNAsine, 0.5  $\mu$ l of SuperScript™ II were added and incubated for 1 hour at  $42^{\circ}\text{C}$ . Then 1  $\mu$ l of 40 mM  $\text{MnCl}_2$  was added and the reaction mixture was incubated for 15 minutes at  $42^{\circ}\text{C}$ . The presence of  $\text{MnCl}_2$  in the reaction mixture allows SuperScript™ to overcome the cap structure during reverse transcription more efficiently: when using 3 mM  $\text{MgCl}_2$  and 2 mM  $\text{MnCl}_2$ , the reverse transcriptase was shown to reveal an extraordinary high cap-dependent transferase activity, and typically the enzyme added preferentially three or four cytosine residues in the presence of 5'-capped mRNA templates (Chenchik *et al.*, 1998, Gene cloning and analysis by RT-PCR, edited by Paul Siebert and James Larrick, BioTechniques Books, Natick, MA; Schmidt and Mueller, 1999, Nucleic Acids Res. 27, 331).

For the PCR reaction, two sets of primers were used for each tested RNA – 3'-specific/5'-specific primers and 3'-specific/tag-specific primers (Fig. 6).

To determine the possibility of using the method of chemical ligation of RNA with tag known sequence specifically to the cap-structure of viral RNAs, the genomic RNA of tobacco mosaic virus (TMV) U1 strain which is known to be capped (Dunigan and Zaitlin, 1990, J. Biol. Chem. 265, 7779-7786.) was used as control. The respective PCR bands were detected when specific primers, U1-Spn and corresponding to RNA-tag primer 779 were used in the PCR reaction (Table 2, Fig. 7).

TABLE 2. Templates and primers used for PCR.

Template	Forward primer	Reverse primer	Corresponding PCR band and detection of cap-structure
Genomic TMV (U1) RNA		U1-Spn	+
Genomic TMV (U1) RNA	779	U1-Spn	+
Non-capped RNA transcript of TMV		U1-Spn	+
Non-capped RNA transcript of TMV	779	U1-Spn	- (non-capped)
Complete cDNA clone of TMV (U1)		U1-Spn	+
Genomic crTMV RNA	K5	2PM	+
Genomic crTMV RNA	779	2PM	+
Non-capped RNA transcript of crTMV	K5	2PM	+
Non-capped RNA transcript of crTMV	779	2PM	- (non-capped)
Complete cDNA clone of crTMV	K5	2PM	+
Subgenomic TMV (U1) RNA for MP	2211	UM50-54	+
Subgenomic TMV (U1) RNA for MP	779	UM50-54	- (non-capped)
Complete cDNA clone of TMV (U1)	2211	UM50-54	+
Subgenomic crTMV RNA for MP	1038	CPF25	+
Subgenomic crTMV RNA for MP	779	CPF25	- (non-capped)
Complete cDNA clone of crTMV	1038	CPF25	0

As a control, the non-capped RNA-transcript of the complete cDNA clone of TMV (U1) was used, and the cap structure was not found as expected (Table 2, Fig. 7).

Then the presence of a cap structure at the 5'-terminus of the genomic RNA of crTMV was tested. For these experiments, the specific PCR primers K5, 2PM and primer 779 which corresponds to the RNA-tag were taken (Table 1, Fig. 7). Interestingly, the mobility of the PCR band observed with the primers 779 and 2PM, was higher than expected (Fig.7). This could reflect the presence of a strong secondary structure at the 5'-terminus of the genomic RNA of crTMV (Dorokhov *et al.*, 1994, FEBS Letters 350, 5-8). This secondary structure is absent at the 5'-terminal part of related TMVs (Goelet *et al.*, 1982, Proc. Natl. Acad. Sci. USA 79, 5818-5822). In control experiments with non-capped transcript of the complete cDNA clone of crTMV, no respective PCR band was observed, as expected.

For subgenomic RNA coding for the TMV (U1) MP gene, the absence of a cap-structure at the 5'-terminus was proposed. We tested the respective sgRNA with the specific primers 2211, UM50-54 and primer 779 corresponding to the RNA-tag. No cap structure was found (Table 2, Fig.7).



The same results were obtained with the respective subgenomic RNA of crTMV (Table 2, Fig. 7) indicating that cap-structure is absent at the 5'-terminus of this subgenomic RNA of tobamoviruses.

Insertion of IRES<sub>MP,75</sub><sup>CR</sup> into a TMV UI based vector that is deficient of MP gene expression, KK6 provides efficient cap-independent MP gene expression.

The KK6 vector (Lehto *et al.*, 1990, *Virology* 174, 145-157) contains two CP subgenomic promoters (sgPr). The first CP sgPr-1 is in its proper place, upstream of the CP gene, whereas the second, CP sgPr-2 is placed upstream of the MP gene. It was shown that the MP gene was expressed via CP sgPr-2 instead of native MP sgPr. As a result of this insertion, KK6 lost the capability of efficient cell-to-cell movement. Analysis showed that I<sub>2</sub> sgRNA does not contain an IRES<sub>MP,75</sub><sup>CR</sup> element in its 5'-nontranslated leader. It has been proposed that IRES<sub>MP,75</sub><sup>CR</sup>-lacking KK6 I<sub>2</sub> sgRNA cannot express the MP gene efficiently. In order to examine this suggestion, IRES<sub>MP,75</sub><sup>CR</sup> was inserted into KK6 between the CP sgPr-2 and the MP gene and we were able to obtain KK6-IRES<sub>MP75</sub> that was stable in progeny (Fig. 8). It was shown that KK6-IRES<sub>MP75</sub> provides synthesis of I<sub>2</sub> sgRNA containing crTMV IRES<sub>MP75</sub> (Fig. 9).

It can be seen that the start of KK6-IRES<sub>MP75</sub> I<sub>2</sub> sgRNA is not changed in comparison to KK6, which means that IRES<sub>MP75</sub> does not serve as MP sgPr.

This insertion drastically improved cell-to-cell movement. KK6 infected Samsun plants systemically but the first symptoms developed slowly (15-17 days) compared to those induced by wild-type TMV (TMV 304) (about 7 days). Symptoms in the upper leaves of KK6-infected plants were distinct: yellow spots in contrast to mosaic symptoms were produced by wild-type TMV.

KK6 virus progeny produced numerous lesions in *N. glutinosa* that developed slower than lesions induced by wild-type TMV UI. The average size of local lesions induced by KK6 was approximately 0.1 mm in comparison to those induced by TMV UI (1.1 mm).

Plants inoculated by KK6-IRES<sub>MP75</sub> looked like KK6-infected Samsun plants but: (i) the first systemic symptoms were developed more rapidly (about 10 days) and (ii) they were much brighter including yellow spots and mosaic. In contrast to KK6 the average size of local lesions induced by K86 in *N. glutinosa* was increased to 0.6-0.7 mm. Examination of the time-course of MP accumulation showed that KK6-IRES<sub>MP75</sub> MP is detected earlier than KK6 MP in inoculated

leaves (Fig. 10). These results allowed the conclusion that insertion of IRES<sub>MP75</sub><sup>CR</sup> upstream of the KK6 MP gene partially restores the movement properties of KK6 defective in cell-to-cell and long-distance transport.

In order to obtain additional evidences of the essential role of IRES in cap-independent MP gene expression of TMV cDNA vectors and in the life cycle of tobamoviruses, series of additional KK6-based vectors was constructed (Fig. 8). KK6-IRES<sub>MP125</sub> contains a natural hairpin-loop structure which is able to inhibit translation of the MP gene *in vitro* in the presence of WT crTMV 5'leader of I<sub>2</sub> sgRNA (Fig. 4) and IRES<sub>MP75</sub>. KK6-H-PL contains a natural hairpin-loop structure and a 72-nt artificial polylinker sequence. KK6-PL contains the polylinker region only. Results of tests for infectivity on *Nicotiana tabacum* cv. Samsun plants (systemic host) are presented in Table.3.

Fig. 11 shows the results of a Western test of CP accumulation in tobacco leaves infected with KK6-based vectors. Replacement of IRES<sub>MP75</sub><sup>CR</sup> by a nonfunctional PL-sequence drastically blocked vector multiplication.

TABLE 3. Virus accumulation in tobacco systemically infected by KK6-based vectors.

cDNA copies	Virus accumulation
TMV 304 (WT)	+++
KK6	+
KK6-IRES <sub>MP75</sub>	++
KK6-IRES <sub>MP125</sub>	++
KK6-H-PL	+/-
KK6-PL	+/-

## EXAMPLE 5

### Creation of artificial, non-natural IRES elements without subgenomic promoter activity provides cap-independent expression of genes of interest in eukaryotic cells

The goal of this example is to demonstrate the approaches for creation of artificial, non-natural IRES elements free of subgenomic promoter activity, which provide cap-independent expression of a gene of interest in eukaryotic cells.

Construction of an artificial, non-natural IRES element on the basis of 18-nt segment of IRES<sub>MP,75</sub><sup>CR</sup>

Analysis of the IRES<sub>MP,75</sub><sup>CR</sup> nucleotide sequence shows that it has a multimer structure and contains four nucleotide sequence segments being a variation of element (-72)GUUUGCUUUUUG(-61) and having high complementarity to *A. thaliana* 18S rRNA (Fig. 12).

In order to design an artificial, non-natural IRES, the 18-nt sequence CGUUUGCUUUUUGUAGUA was selected.

Four oligos were synthesized:

MP1(+):

5'-CGCGCAAGCTTTGCTTTTTGTAGTACGTTTGCTTTTTGTAGTACTGCAGGCGGG -3'

MP1(-):

5'-CCCGCCTGCAGTACTACAAAAGCAAACGTACTACAAAAGCAAAGCTTGCGCG - 3'

MP2(+):

5'-GGCGGCTGCAGTTTGCTTTTTGTAGTACGTTTGCTTTTTGTAGTAGAATTCGG-GC-3'

MP2(-):

5'-GCCCGAATTCTACTACAAAAGCAAACGTACTACAAAAGCAAAGTGCAGCCG-CC-3'

Primers MP1(+) and MP1(-) were annealed to each other yielding dsDNA fragment A:

CGCGCAAGCTTTGCTTTTTGTAGTACGTTTGCTTTTTGTAGTACTGCAGGCGGG  
GCGCGTTGAAACGAAAAACATCATGCAAACGAAAAACATCATGACGTCCGCCC

HindIII

PstI

Primers MP2(+) and MP2(-) were annealed to each other yielding dsDNA fragment B:

GGCGGCTGCAGTTTGCTTTTTGTAGTACGTTTGCTTTTTGTAGTAGAATTCGGGC  
CCGCCGACGTCAAACGAAAAACATCATGCAAACGAAAAACATCATCTTAAGCCCG

PstI

EcoRI

Both fragments were digested with PstI and ligated to each other. Then the ligation product A+B was extracted using agarose electrophoresis and digested with HindIII and EcoRI followed

by ligation into the hGFP-GUS vector described by Skulachev *et al.* (1999, Virology 263, 139-154) using HindIII and EcoRI cloning sites (Fig. 13).

## Results

The transcripts depicted in Fig. 13 were translated in rabbit reticulocyte lysate (RRL) as described by Skulachev *et al.* (1999, Virology 263, 139-154) and synthesized products were analyzed by gel electrophoresis. Results represented in Fig. 13 show that an artificial, non-natural sequence based on a 18-nt segment of IRES<sub>MP,75</sub><sup>CR</sup> provides 3'-proximal-located GUS gene expression. This means that two features, namely complementarity to 18S rRNA and multimer structure are essential for IRES<sub>MP,75</sub><sup>CR</sup> function and effectiveness.

A tetramer of 18-nt segment does not reach the level of IRES<sub>MP,75</sub><sup>CR</sup> activity but there is a way to improve the activity of artificial, non-natural IRES elements using the 12-nt segment GCUUGCUUUGAG which is complementary to 18S rRNA.

### Construction of an artificial, non-natural IRES using 19-nt segment of IRES<sub>CP,148</sub><sup>CR</sup>

Analysis of structural elements essential for IRES<sub>CP,148</sub><sup>CR</sup> activity (Figs. 14-17) shows that a polypurine (PP) segment is crucial for IRES<sub>CP,148</sub><sup>CR</sup> functioning. As a prominent element of the PP tract, a 9-nt direct repeat in 19-nt sequence: AAAAGAAGGAAAAAGAAGG (called direct repeat (DR)) was used for the construction of an artificial IRES. In order to obtain the tetramer of DR the following primers were used:

CP1(+):

5'-CGCGCAAGCTTAAAAGAAGGAAAAAGAAGGAAAAGAAGGAAAAAGAAGGCT-GCAGGCGGG-3'

CP1(-):

5'-CCCGCCTGCAGCCTTCTTTTCCTTCTTTTCCTTCTTTTCCTTCTTTTAAGCT-TGCGCG-3'

CP2(+):

5'-GGCGGCTGCAGAAAAGAAGGAAAAAGAAGGAAAAGAAGGAAAAAGAAGGAA-TTCGGGC-3'

CP2(-):

5'-GCCCGAATTCCTTCTTTTCCTTCTTTTCCTTCTTTTCCTTCTTTTCTGCAGC-CGCC-3'

According to the experimental procedure described above, the following IRES element was used as intercistronic spacer:

5'-CGCGCAAGCUUAAAAGAAGGAAAAAGAAGGAAAAGAAGGAAAAAGAAGGCU-GCAG  
AAAAGAAGGAAAAAGAAGGAAAAGAAGGAAAAGAAGGAAUUCAUG-3'

### Results

The transcripts depicted in Fig. 13 were translated in rabbit reticulocyte lysate (RRL) as described by Skulachev *et al.* (1999, Virology 263, 139-154) and synthesized products were analyzed by gel electrophoresis. The results represented in Fig. 13 show that an artificial, non-natural sequence based on repeated 19-nt segment of IRES<sub>CP,148</sub><sup>CR</sup> provides the efficient expression of a 3'-proximally located GUS gene.

### EXAMPLE 6

#### TMV cDNA transcription vector expressing a replicase gene in infected cells cap-independently

The main goal of this example was to obtain two new TMV U1-based viruses with modified 5'UTR providing expression of the replicase gene in a cap-independent manner:

- 1) Omega-leader of TMV was completely substituted by IRES<sub>MP,75</sub><sup>CR</sup>.

GUUCGUUUCGUUUUUGUAGUAUAAUUAUUUUGUCAGAUAGAGAUUGGUUAGAG  
AUUUGUUCUUUGUUUGACCAUGG.

- 2) Since it is believed that the first 8 nucleotides of the TMV 5'UTR are essential for virus replication (Watanabe *et al.*, 1996, J. Gen. Virol. 77, 2353-2357), IRES<sub>MP,75</sub><sup>CR</sup> was inserted into TMV leaving the first 8 nucleotides intact:

GUAUUUUUGUAGUAUAAUUAUUUUGUCAGAUAGAGAUUGGUUAGAGAUUUGUU  
CUUUGUUUGACCAUGG.

The following primers were used:

- a) SP6-IRES-1 (in the case of the first variant)

XbaI                  SP6 Promotor                  IRES<sub>MP,75</sub><sup>CR</sup>  
GGGTCTAGATTTAGGTGACACTATAGTTCGTTTCGTTTTGTAGTA

b) SP6-IRES-2 (in the case of the second variant)

XbaI                  SP6 Promotor                  IRES<sub>MP,75</sub><sup>CR</sup>

GGGTCTAGATTAGGTGACACTATAGTATTTTGTAGTATAATTAAATATTTGTC.

c) IRES-NcoI (reverse primer to obtain IRES with a NcoI site at 3'end):

GGGCCATGGTCAAACAAAGAACAAATCTCTAAAC.

d) TMV-NcoI (direct primer to obtain TMV polymerase, starting from NcoI site):

NcoI

GGGCCATGGCATACACACAGACAGCTAC.

e) TMV-Xho (reverse primer to obtain 5'-part of replicase from AUG to SphI site)

XhoI

ATGTCTCGAGCGTCCAGGTTGGGC.

Cloning strategy:

PCR fragment A was obtained using oligos SP6-IRES1 and IRES-NcoI and crTMV clone as template. PCR fragment B was obtained using oligos TMV-NcoI and TMV-XhoI and TMV-304L clone. Fragments A and B were cloned simultaneously into the pBluscriptSK+ vector using XbaI and XhoI sites (fragments were ligated together through NcoI site). The same procedure was applied to obtain the second variant of the virus using SP6-IRES2 oligo.

At the next stage, the whole TMV cDNA was cloned into the obtained vector using SphI and KpnI sites to restore the viral genome (Fig. 18).

## EXAMPLE 7

Tobamoviral vectors Act2/crTMV and Act2/crTMV IRES<sub>MP,75</sub><sup>CR</sup>-GUS based on Actin 2 transcription promoters

The main goal of this example is the demonstration of the construction strategy of a new crTMV-based vector with which viral genome expression in plant cells occurs under the control of an efficient Actin 2 transcription promoter. It allows the use of the vector Act2/crTMV/ IRES<sub>MP,75</sub><sup>CR</sup>-GUS for gene expression in plants.

#### Cloning Act2 into pUC19

The Act2 transcription promoter (about 1 000 bp) was cut out of plasmid pACRS029 by digestion with KpnI and PstI and cloned into pUC19 digested with KpnI and PstI.

#### Creation of a PstI site in plasmid T7/crTMV (see Fig. 1) upstream of crTMV genome start

334-nt cDNA fragment of the 5'-terminal portion of the crTMV genome obtained by PCR using the direct primer ATGCTGCAGGTTTTAGTTTTATTGCAACAACAA (the PstI site is underlined) and the reverse primer ATGCGATCGAAGCCACCGGCCAAGGAGTGCA (PvuI site is also underlined) was digested with MunI and PstI and inserted into T7/crTMV between KpnI and MunI restriction sites together with the Actin2 promoter (KpnI-PstI fragment from pUCAct2).

#### Fusion of 5'-terminus of crTMV to Act2 transcriptional start without additional sequences

This step was carried out by site-directed mutagenesis using oligonucleotide primer specific for both Act2 and crTMV to obtain the final construct Act2/crTMV (FIGURE 19).

To get the vector Act2/crTMV/ IRES<sub>MP,75</sub><sup>CR</sup>-GUS (Fig. 20) the XhoI-NotI cDNA fragment of plasmid Act2/crTMV (FIGURE 19) was replaced by the XhoI-NotI DNA fragment of T7/crTMV/ IRES<sub>MP,75</sub><sup>CR</sup>-GUS construct (Fig. 2) that contains the GUS gene under the control of IRES<sub>MP,75</sub><sup>CR</sup>.

### EXAMPLE 8

#### Construction of circular single-stranded tobamoviral vector KS/Act2/crTMV/IRES<sub>MP,75</sub><sup>CR</sup>-GUS (Fig. 21)

The main goal of this example is to demonstrate the possibility of using circular single-stranded DNA vectors for foreign gene expression in plants.

In order to construct KS/crTMV/IRES<sub>MP,75</sub><sup>CR</sup>-GUS (Fig. 21), 9.2 kb KpnI-NotI cDNA

fragment of vector Act2/crTMV/IRES<sub>MP,75</sub><sup>CR</sup>-GUS was inserted into plasmid pBluescript II KS+ (Stratagene) digested with KpnI-NotI and containing the phage f1 replication origin. Single-stranded DNA of vector KS/Act2/crTMV/IRES<sub>MP,75</sub><sup>CR</sup>-GUS was prepared according to Sambrook *et al.*, 1989 (Molecular Cloning: a Laboratory Manual, 2ed edn. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York) and used in particle bombardment experiments with *Nicotiana benthamiana* leaves (see previous example). GUS expression was detected by usual histochemical staining 2-3 days after shooting.

## EXAMPLE 9

### Construction of tobamoviral vector KS/Act2/crTMV-Int/IRES<sub>MP,75</sub><sup>CR</sup>-GUS containing oleosin intron from *Arabidopsis thaliana*

The main goal of this example is to create vector KS/Act2/crTMV/IRES<sub>MP,75</sub><sup>CR</sup>-GUS containing *Arabidopsis thaliana* oleosin gene intron that should be removed after transcript processing (Fig. 22).

The cloning strategy comprised the following steps:

#### 1. Cloning of *A. thaliana* oleosin gene intron.

*A. thaliana* oleosin gene intron was obtained by PCR using *A. thaliana* genomic DNA and specific primers : A.th./Int (direct) ATGCTGCAGgttttagttCAGTAAGCACACATTTATCATC (PstI site is underlined, lowercase letters depict crTMV 5'terminal sequence) and A.th./Int (reverse) ATGAGGCCTGGTGCTCTCCCGTTGCGTACCTA (StuI is underlined).

#### 2. Insertion of *A. thaliana* oleosin gene intron into 334-nt 5'-terminal fragment of crTMV cDNA.

cDNA containing *A. thaliana* oleosin gene intron was digested with PstI/StuI and ligated with DNA fragment obtained by PCR using primers corresponding to positions 10-334 of crTMV genome: atgAGGCCTTTATTGCAACAACAACAACAATTA (StuI site is underlined) and ATGCGATCGAAGCCACCGGCCAAGGAGTGCA (PvuI site is underlined).

The next steps were as described in example 7 (see also example 18).



**EXAMPLE 10**

Influence of rapamycin as an inhibitor of cap-dependent initiation of translation on GUS gene expression in tobacco protoplasts transfected with IRES<sub>MP,75</sub><sup>CR</sup> containing bicistronic transcription vectors, 35S/CP/IRES<sub>MP,75</sub><sup>CR</sup>/GUS (Fig. 23) and 35S/GUS/ IRES<sub>MP,75</sub><sup>CR</sup>/CP (Fig. 24)

The aim of this example is to demonstrate the principal possibility to use inhibitors of cap-dependent translation to increase efficiency of IRES-mediated cap-independent translation of a gene of interest.

Rapamycin as an inhibitor of cap-dependent initiation of translation was selected. Recently, a novel repressor of cap-mediated translation, termed 4E-BP1 (eIF-4E binding protein-1) or PHAS-1 was characterized (Lin *et al.*, 1994, Science 266, 653-656; Pause *et al.*, Nature 371, 762-767). 4E-BP1 is a heat- and acid-stable protein and its activity is regulated by phosphorylation (Lin *et al.*, 1994 Science 266, 653-656; Pause *et al.*, Nature 371, 762-767). Interaction of 4EBP1 with eIF-4E results in specific inhibition of cap-dependent translation, both *in vitro* and *in vivo* (Pause *et al.*, Nature 371, 762-767). It has been shown that rapamycin induces dephosphorylation and consequent activation of 4E-BP1 (Beretta *et al.*, 1996, EMBO J. 15, 658-664).

Construction of IRES- and GUS gene-containing vectors 35S/CP/ IRES<sub>MP,75</sub><sup>CR</sup>/GUS (Fig. 23), 35S/GUS/ IRES<sub>MP,75</sub><sup>CR</sup>/CP (Fig. 24) and a method of tobacco protoplast transfection with 35S-based cDNA were described by Skulachev *et al.* (1999, Virology 263, 139-154). Comparison of GUS gene expression in tobacco protoplasts treated by rapamycin and transfected with bicistronic cDNA with GUS gene in 3'- and 5'-proximal location shows the possibility to increase IRES-mediated cap-independent translation of the GUS gene.

**EXAMPLE 11**

Influence of potyvirus VPg as a inhibitor of cap-dependent initiation of translation on GUS gene in tobacco protoplasts transfected with IRES<sub>MP,75</sub><sup>CR</sup> containing bicistronic transcription vectors 35S/CP/IRES<sub>MP,75</sub><sup>CR</sup>/GUS (Fig. 23) and 35S/CP-VPg/ IRES<sub>MP,75</sub><sup>CR</sup>/GUS

This example demonstrates the principal possibility of using a gene product to inhibit cap-dependent translation (Fig. 25). Recently, interaction between the viral protein linked to the genome (VPg) of turnip mosaic potyvirus (TuMV) and the eukaryotic translation initiation

factor eIF(iso)4E of *Arabidopsis thaliana* has been reported (Wittman *et al.*, 1997, Virology 234, 84-92). Interaction domain of VPg was mapped to a stretch of 35 amino acids and substitution of an aspartic acid residue within this region completely abolished the interaction. The cap structure analogue m<sup>7</sup>GTP, but not GTP, inhibited VPg-eIF(iso)4E complex formation, suggesting that VPg and cellular mRNAs compete for eIF(iso)4E binding (Leonard *et al.*, 2000, J. Virology 74, 7730-7737).

The capability of VPg to bind eIF(iso)4E could be used for inhibition of cap-dependent translation. We propose to use the vector 35S/CP-VPg/IRES<sub>MP,75</sub><sup>CR</sup>/GUS (Fig. 25) wherein CP is fused with VPg from potyvirus potato virus A. Comparison of GUS gene expression in protoplasts transfected with 35S/CP-VPg/IRES<sub>MP,75</sub><sup>CR</sup>/GUS or 35S/CP/IRES<sub>MP,75</sub><sup>CR</sup>/GUS would allow to increase IRES-mediated and cap-independent GUS gene expression.

## EXAMPLE 12

### In vivo genetic selection of an IRES sequence or a subgenomic promoter using TMV vector

This example demonstrates the possibility of using *in vivo* genetic selection or Systematic Evolution of Ligands by Exponential enrichment (SELEX) of a subgenomic promoter or an IRES sequence providing cap-independent expression of a gene of interest in a viral vector. This approach proposes using side-by-side selection from a large number of random sequences as well as sequence evolution (Ellington and Szostak, 1990, Nature 346, 818-822; Tuerk and Gold, 1990, Science 249, 505-510; Carpenter and Simon, 1998, Nucleic Acids Res. 26, 2426-2432).

The project encompasses:

1. *In vitro* synthesis of crTMV-based defective-interfering (DI) transcript containing the following elements (5'-3' direction): (i) a T7 transcription promoter, (ii) a 5'-terminal part of crTMV genome with a sequence responsible for viral genome complementary (minus chain) synthesis, (iii) a sequence coding for the N-terminal part of a viral replicase, (iv) a sequence containing 75-nt randomized bases, (v) a neomycin phosphotransferase II (NPT II) gene, (vi) a crTMV origin of assembly (Oa), and (vii) a 3'-terminal part of the crTMV genome with minus chain genome promoter sequence (Fig. 26).
2. Co-transfection of tobacco protoplasts by a transcript together with crTMV genomic RNA (Fig. 1). Protoplasts will grow and regenerate in media containing kanamycin.

3. Selection and isolation of an IRES or a subgenomic promoter element providing protoplast survival and regeneration in the presence of kanamycin.

### EXAMPLE 13

#### Construction of TMV-U1-based vector containing heterologous viral IRES

The crTMV-based set of vectors described in example 2 contains homologous viral IRES sequences taken from the same crTMV genome. This creates direct repeat of different length, which quite often causes instability of the vector during plant infection (see Chapman *et al.*, 1992, Plant J. 2(4), 549-557, Shivprasad *et al.*, 1999, Virology 255, 312-323). To avoid that, the combination of TMV-U1 genome and heterologous IRES<sub>MP,75</sub><sup>CR</sup> sequence was chosen. Another reason to try a different tobamovirus for vector construction is that - in contrast to crTMV - TMV-U1 has a more limited host range (see table 1), but is also more virulent in *Nicotiana* species, for example it accumulates to a higher level and shows more severe symptoms in *N.benthamiana* and *N.tabacum*.

Plasmid TMV304 (fig. 27) (Dawson *et al.*, 1986, Proc. Natl. Acad. Sci. USA 83, 1832-1836; Lehto *et al.*, 1990, Virology 174, 145-157) was taken as the starting material. Four primers were ordered to introduce additional HindIII and XbaI restriction sites into the viral genome:

#### 1.TMVvect1Nco

5'- acggagggcccatggaacttaca - 3'

#### 2.TMVvect2Hind

5'- ctagaagcttcaagttgcaggaccagaggtccaaa - 3'

#### 3.TMVvect3Xba

5'- ctagtctagaggtagtcgaagatgcataataaataac - 3'

#### 4.TMVvect4Kpn

5'- gtacgggtacctgggcccctaccgggggtaacgggggggattc - 3'.

Oligonucleotides 1 and 2 were used for PCR amplification of the CP and C-terminal part of the MP genes, 3 and 4 – to amplify the 3'-nontranslated region (fig.28). Then both PCR products were digested with NcoI/HindIII or XbaI/KpnI and cloned into TMV304 between unique NcoI and KpnI restriction sites together with either IRES<sub>MP,75</sub><sup>CR</sup>-GUS insert (taken from pICH766, HindIII/XbaI) or IRES<sub>MP,75</sub><sup>CR</sup>-eGFP (pICH1041, HindIII/XbaI) insert (four-fragment ligation) (fig.28). As a result, constructs pICH1865 (with GFP) and pICH1871 (with GUS) were obtained (fig.29). For plant infection, these plasmids were linearized with KpnI, transcribed *in vitro* using SP6 promoter and inoculated onto *N.benthamiana* plants as described previously. GUS-staining was performed 7 days post inoculation (dpi) (see example 3). Fig. 31A shows GUS expression in the inoculated, but not in the systemic leaves. Similar results were obtained with GFP-containing viral constructs.

#### EXAMPLE 14

TMV-U1-based vector: a foreign gene can be expressed via an IRES of plant origin or via a synthetic IRES that are free of subgenomic promoter activity.

Two additional TMV-U1-based vectors were constructed. Different non-viral IRES sequences were used for cloning: firstly, the 453-nt 5'-nontranslated leader sequence of *Nicotiana tabacum* heat shock factor 1 (NtHSF-1, EMBL/Genbank nucleotide database, accession number ABO14483) and, secondly, artificial sequence (GAAA)x16. Both sequences showed IRES activity *in vitro* (rabbit reticulocyte lysate, wheat germ extract) and *in vivo* (tobacco protoplasts, HeLa cells).

To get the new versions of TMV-U1 based vector, pICH1871 (TMV-U1-GUS) plasmid (see fig.29 and the previous example) was digested with NcoI and SalI and ligated with two inserts: NcoI/HindIII fragment (from the same construct, CP and partially MP gene) and HindIII/SalI fragments (IRES-GUS) from the plasmids hGFP-NtHSF-GUS and hGFP-(GAAA)x16-GUS (unpublished) (fig.30).

The inoculation of transcripts obtained from pICH4235 (with NtHSF sequence) and pICH4246 (with artificial IRES GAAAx16) onto *N. benthamiana* plants was performed in an usual way (see previous examples). GUS expression was analysed 7dpi. The results are showing that in both cases the expression level of GUS gene is comparable to that achieved by an IRESs of viral origin (for example, IRES<sub>MP,75</sub><sup>CR</sup> or IRES<sub>CP,148</sub><sup>CR</sup>, fig. 31 B, C). It is clear that IRES sequences used in those constructs (taken from the plant genome or created artificially)

are definitely free of any subgenomic promoter activity. Infection with pICH4246 also showed instability of the vector – like all the other viral constructs containing the GUS gene, this one reverted to wild type and symptoms of systemic infection appeared quite soon (7-8 dpi). In case of pIC4235 (NtHSF sequence) – symptoms in the upper leaves were not visible even 14 dpi and started to appear only 20-21 dpi. This means that 4235, which carries a long (453 b.p.) and highly structured non-viral IRES sequence is much more stable than the other related vectors and gives a good chance for stable systemic expression of genes which are smaller than GUS (1.8 kb), for example GFP.

### EXAMPLE 15

Agroinfiltration provides a rapid, cheap and efficient method to express the foreign proteins via IRES-based viral vectors in plants.

As the first step, Act2/crTMV/IRES<sub>CP,148</sub><sup>CR</sup>-GFP construct (plasmid pICH3011), was cloned into the binary vector pICBV10 (Icon Genetics GmbH). pICBV10 was digested by KpnI and HindIII and ligated with KpnI/NotI fragment from pICH3011 and nos transcriptional terminator (NotI/HindIII fragment from the same construct). The resulting plasmid (pICH4471) was transformed into *Agrobacterium tumefaciens* (strain GV3101). Colonies were grown overnight in a 5 ml of a liquid culture and agroinfiltration into *Nicotiana benthamiana* plants was performed using a common procedure. GFP expression in the inoculated leaves was detectable with the UV lamp 6-7 days after infiltration.

### EXAMPLE 16

Expression of pharmaceutical proteins from the tobamoviral vector Act2/crTMV/IRES<sub>CP,148</sub><sup>CR</sup>

For pharmaceutical protein expression in plant leaves, crTMV-based viral vector under the control of *Arabidopsis* actin 2 promoter was used (An et al., 1996, Plant J. 10, 107-121). This basic vector constructed pIC3011 is able to express via internal translation initiation the foreign genes (for example, GFP) inserted downstream of IRES<sub>CP148</sub>. To express the Hepatitis B protein in plants, corresponded crTMV-based viral vector was constructed. The Hepatitis B protein gene was inserted into pIC3011 subsequently the additional IRES<sub>CP148</sub> placed between CP gene and 3'-terminal nontranslated viral sequence. Resulting plasmid was designated pICP1260 (#62C, Arab. Act2promoter: crTMV: IRES<sub>CP148</sub>cr: hepatitis B protein). It

was precipitated with tungsten particles and used for bombardment experiments. Particle bombardment of detached *Nicotiana benthamiana* leaves was performed using the flying disk method with a high-pressure helium-based PDS-1000 apparatus (Bio-Rad) as described in Morozov *et al.* (1997, Journal of General Virology **78**, 2077-2083). pICP1260-bombarded *N.benthamiana* leaves were tested by Western blotting 4 days post bombardment (d.p.b.) and showed some expression of hepatitis B protein (less than 0,05% of total protein).

For the expression of human antibodies (FAT and OAT heavy and light chains; received from Sunol), another pIC3011-based vectors were constructed. Heavy chains of humanized anti-TF Mega IgG1 (FAT) and IgG4 (OAT) fused with plant signal peptide were cloned into crTMV Arab. Act2-driven vector to give pICP1284 (#101C, Arab. Act2promoter: crTMV: IREScp148(cr): pspFAT-HC) and pICP1283 (#89C, Arab. Act2promoter: crTMV: IREScp148(cr): pspOAT-HC). Light chain pspLCIgGE:E was fused with plant signal peptide and cloned into crTMV Arab. Act2-driven vector to give pICP1288 (#208C, Arab. Act2promoter: crTMV: IREScp148(cr): pspLCIgGE:E). Then HCs and LC coding constructs were bombarded into detached *N.benthamiana* leaves. Additionally the ratio between HCs and LC-expressed constructs was varied in co-bombardment (1:1, 2:1, 3:1), and the bombarded leaves are tested 5, 6, 7, 8 days post bombardment. ELISA for assembled IgG and Western blots showed a generation of well measurable amounts of protein (both heavy and light chain fragments). However the significant over-expression of LC compared to HCs was detected by Western blots. The best expression was found to be 7 d.p.b. with the ratio HC/LC 1:2 (data not shown).

### Example 17

#### Construction of a TMV cDNA transcription vector expressing a replicase gene in infected cells in a cap-independent manner

The main goal of this example was to obtain six new TMV U1-based viruses with modified 5'UTR providing expression of the replicase gene in a cap-independent manner (parts of TMV-U1 omega sequence are underlined):

1) Control mutant of the wild-type TMV-U1 – NcoI site is introduced at the initiation codon of the replicase gene:

GUAUUUUUACAACAAUUACCAACAACAACAAACAACAACAAUUAACAUAUACUAU

UUACAATTACCAUGG

2) Omega-leader of TMV-U1 was completely substituted by IRES<sub>MP,75</sub><sup>CR</sup>:

GUUCGUUUCGUUUUUGUAGUAUAAUUAUUUUGUCAGAUAGAGAUUGGUUAGAG  
AUUUGUUCUUUGUUUGACCAUGG.

3-4) Since it is believed that the first 8 nucleotides of the TMV-U1 5'UTR are essential for virus replication (Watanabe *et al.*, 1996, J. Gen. Virol. 77, 2353-2357), IRES<sub>MP,75</sub><sup>CR</sup> was inserted instead of the TMV-U1 omega leaving either the first 8 nucleotides intact:

GUAUUUUUUUCGUUUGCUUUUUGUAGUAUAAUUAUUUUGUCAGAUAGAGAUUG  
 GUUAGAGAUUUGUUCUUUGUUUGACCAUGG,

or the first 18 nucleotides intact:

GUAUUUUUACAACAAUUAUUCGUUUGCUUUUUGUAGUAUAAUUAUUUUGUCAGA  
 UAAGAGAUUGGUUAGAGAUUUGUUCUUUGUUUGACCAUGG.

5) IRES<sub>MP,75</sub><sup>CR</sup> was inserted between nucleotides 8 and 18 of the omega leader:

GUAUUUUUUUCGUUUGCUUUUUGUAGUAUAAUUAUUUUGUCAGAUAGAGAUUG  
 GUUAGAGAUUUGUUCUUUGUUUGACCAACAACAACAACAACAACAACAACA  
UACUAUUUACAATTACCAUGG,

6) IRES<sub>MP,75</sub><sup>CR</sup> was inserted between nucleotides 18 and 19 of the omega leader:

GUAUUUUUACAACAAUUAUUCGUUUGCUUUUUGUAGUAUAAUUAUUUUGUCAGA  
 UAAGAGAUUGGUUAGAGAUUUGUUCUUUGUUUGACCAACAACAACAACAACAACA  
ACAUUACAACUACUAUUUACAATTACCAUGG

The following primers were used:

1) H3-T7-omega (in the case of the first variant):

HindIII    T7 Promoter                      omega

5'- ctagaagcttaatacgcactatagtagtattttacaacaattaccaacaac - 3'

2) H3-T7-IRESmp (in the case of the second variant):

51

HindIII T7Promoter

IRES<sub>MP,75</sub><sup>CR</sup>5'- ctagaagcttaatacgactcactatagttcgttgctttttagtataattaaa - 3'

3) H3-T7-8U1-IRESmp (in the case of the third variant):

HindIII T7Promoter

omega/IRES<sub>MP,75</sub><sup>CR</sup>5'- ctagaagcttaatacgactcactatagttcgttgctttttagtataattaaa - 3'

4) H3-T7-18U1-IRESmp (in the case of the fourth variant):

HindIII T7Promoter

omega/IRES<sub>MP,75</sub><sup>CR</sup>5'- ctagaagcttaatacgactcactatagttttacaacaattatcgttgctttttagtataattaaa - 3'

For the omega versions 5 and 6 two more oligonucleotides were ordered in addition to primers 3 and 4:

5) IRESmp-19U1-plus:

IRES<sub>MP,75</sub><sup>CR</sup>/omega

5'- gtttagagattgttcttgttggatccaacaacaacaacaacaacatt - 3'

6) 19U1-IRESmp-minus:

IRES<sub>MP,75</sub><sup>CR</sup>/omega

5'- aatgtgtgttgttgttgttggatcaaacaagaacaaatctctaaac - 3'

The rest of the primers that were used to obtain the omega mutants:

7) IRESmp-NcoI (reverse primer to obtain IRES with the NcoI site at 3'end):

5'-gggccatgggtcaaacaagaacaaatctctaa-3'.



8) U1-Repl-Nco-plus (direct primer to obtain TMV-U1 polymerase, starting from the NcoI site):

5'- cgtacatggcatacacacagacagctaccacatca - 3'

9) U1-Repl-Sph-minus (reverse primer to obtain 5'-part of replicase from AUG to SphI site)

5'- tccaggttgggcatgcagcagtgtag - 3'

10) Omega-Nco-minus (reverse primer to obtain 3'end of the omega sequence with the NcoI site at the replicase AUG codon)

5'- cgtacatggtaattgtaaataagtaattgtaag - 3'

#### Cloning strategy:

TMV304 clone (fig. 27) (Dawson *et al.*, 1986, Proc. Natl. Acad. Sci. USA 83, 1832-1836; Lehto *et al.*, 1990, Virology 174, 145-157) served as a template for all the PCR reactions with omega-specific primers; IRES<sub>MP,75</sub><sup>CR</sup> was amplified from the plasmid pICH766.

PCR fragment 1 was obtained using primers 1 and 10; fragment 2 with primers 2 and 7. For the fragments 3 and 4 oligonucleotide combinations 3+7 and 4+7 were used.

PCR fragments 5 and 6 were amplified in two steps. Firstly, intermediate fragments 5a, 6a (primers 3+6 and 4+6) and 5b (5+10) were obtained. Then fragments 5a/5b and 6a/5b were annealed to each other and used for amplification with the following primer combinations: 3+10 and 4+10 to get the final PCR products 5 and 6. N-terminal part of the TMV-U1 replicase (PCR fragment 7, nucleotide positions in the genome 68-450) was amplified with the oligonucleotides 8 and 9 to introduce NcoI site at the beginning of the replicase gene. Fragment 1 together with the fragment 7 was cloned simultaneously into the pUC19 vector using HindIII and SphI sites (fragments were ligated through the NcoI site, resulting plasmid pICH4552). The same cloning procedure was applied to obtain all the other variants (PCR products 2-6) of the intermediate construct (HindIII-T7promoter-omega mutant-NcoI-Replicase-SphI, plasmids pICH4565, pICH4579, pICH4584, pICH4597, pICH4602).

At the next stage HindIII/SphI fragment from each of the intermediate constructs was cloned together with the EheI/HindIII fragment from pUC18 into the TMV304 plasmid (see fig. 27) between Ehe I and SphI restriction sites to obtain the final full-length cDNA constructs of

the TMV-U1 mutants (6 different versions of the omega region with and without IRES<sub>crmp75</sub>, plasmids pICH 4735, 4744, 4752, 4765, 4771, 4788).

These constructs were transcribed *in vitro* and tested for infectivity on *Nicotiana benthamiana* (systemic host) and *Nicotiana tabacum* Samsun NN (necrotic host) together with the TMV-U1 wild-type clone (TMV304). Wild-type virus and pICH 4735 (control mutant, NcoI site is introduced at the beginning of the replicase gene) were showing systemic infection on *N. benthamiana* and typical necrotic lesions on Samsun NN plants 3-4 days post inoculation (dpi). None of the other mutants caused local lesions on the NN plants, but at least one construct pICH 4771 (omega 1-8 b.p./IRES<sub>mp75</sub>/omega 18-67 b.p.) caused clear symptoms of systemic spread; development of these symptoms was delayed comparing to the wild-type TMV-U1 and pICH 4735 infection (7dpi). This result shows the principal possibility to express the viral replicase gene in a cap-independent manner, for example, to infect the plants with the uncapped RNA transcripts which might be translated from IRES<sub>MP,75</sub><sup>CR</sup> or any other known IRES that is functional in a plant cell.

### Example 18

#### Construction of tobamoviral vectors Act2/crTMV and Act2/crTMV IRES<sub>MP,75</sub><sup>CR</sup>(IRES<sub>CP,148</sub><sup>CR</sup>)-GFP based on Actin 2 transcription promoters

The main goal of this example is the demonstration of the construction strategy of a new crTMV-based vector with which viral genome expression in plant cells occurs under the control of an efficient Actin 2 transcription promoter from *Arabidopsis thaliana* (An *et al.*, 1996, Plant J., 10, 107-121). It allows the use of the vectors Act2/crTMV/IRES<sub>MP,75</sub><sup>CR</sup>-GFP and Act2/crTMV/IRES<sub>CP,148</sub><sup>CR</sup>-GFP for gene expression in plants.

#### 1. Act2 promoter cloning into pUC19

The Act2 transcription promoter was cut out of plasmid pACRS029 (pIC04) by digestion with KpnI and Pst and cloned into pUC19 digested with KpnI and PstI (construct pICH1364).

#### 2. Fusion of the 5'-terminus of crTMV genome to Act2 transcriptional start without additional sequences.

For this step, the following primers were used:

## 1) BsrGI-Act2:

5'-ccattatttaagtacataactaatcgt- 3'

## 2) PvuI-cr:

5'-tccaactcaagcgatcgaaagcca- 3'

## 3) Act2-cr-plus:

5'-catatattttcctctccgcttgaagtttttagttttattgcaacaacaac- 3'

## 4) cr-Act2-minus:

5'-gttggtgttgcaataaaaactaaaactcaaagcggagaggaaaatatatga- 3'.

PCR fragment 1 was obtained with primers 1 and 4, fragment 2 was amplified using oligonucleotides 2 and 3. Then both fragments were annealed to each other and used for the second round of amplification with the primers 1 and 2 to get the PCR product 3, which was cloned into pGEM-T vector (Promega). As a result, construct pICH1823 that contains 3'-end of the Actin2 promoter (from BsrGI site to transcription start) and the 5'-terminal part part of the crTMV genome (until the unique PvuI site) was obtained. In this construct the first nucleotide of the viral genome (G) was located immediately downstream of the proposed transcriptional start (A) of the Actin2 promoter, so the expected viral-specific transcript should contain one additional nucleotide (A) at the 5'-end, which is usually not affecting the efficient replication of the viral genome.

## 3. Cloning of the rest of the genome together with the last construct.

Construct pICH1364 was digested with BsrGI/HindIII and ligated together with the the following fragments: BsrGI/PvuI from pICH1823, PvuI/SacI and SacI/BamHI taken from the crTMV cDNA clone and BamHI/HindIII insert from the plasmid pIC02 (nos transcriptional terminator). The final construct (pICH1983) was tested in particle bombardment experiments

with *Nicotiana benthamiana* leaves as described previously (Morozov *et al.*, 1997, Journal of General Virology 78, 2077-2083) and the infectivity was checked after reinoculation of *Nicotiana tabacum* Samsun NN plants (necrotic host) with the *N.benthamiana* leaf material 3 days after bombardment.

#### 4. Cloning of the vectors with Actin2 promoter containing GUS and GFP genes.

To get the final vector constructs, XhoI/NotI fragments from either T7/crTMV/IRES<sub>MP,75</sub><sup>CR</sup>-GUS and T7/crTMV/IRES<sub>CP,148</sub><sup>CR</sup>-GUS (Fig. 2) or T7/crTMV/IRES<sub>MP,75</sub><sup>CR</sup>-GFP and T7/crTMV/IRES<sub>CP,148</sub><sup>CR</sup>-GFP were cloned into the pIC1823 construct. The resulting plasmids were also tested by particle bombardment and showed GUS and GFP expression in the *Nicotiana benthamiana* leaves.

**Claims**

1. Vector capable of amplification and expression of a gene in a plant comprising a nucleic acid having a sequence for at least one non-viral gene to be expressed and having or coding for at least one IRES element necessary for translation of a gene downstream thereof.
2. Vector according to claim 1 wherein the IRES element is located upstream of said non-viral gene to be expressed for directly supporting its translation.
3. Vector according to one of claims 1 or 2 wherein the IRES element indirectly supports the translation of the non-viral gene to be expressed by directly supporting the translation of another gene downstream thereof which is essential for a function of said vector selected from the group of infection, amplification, virus assembly, ability to suppress the silencing of viral infection development in plant cells, ability to redirect the metabolism in plant cells, and cell-to-cell or long-distance movement of said vector.
4. Vector according to one of claims 1 to 3 further comprising at least a portion of a sequence of the host plant genome in an anti-sense orientation for suppressing a gene of the host plant.
5. Vector according to claim 4 wherein said sequence in anti-sense orientation suppresses a gene essential for cap-dependent translation in plants.
6. Vector capable of amplification in a plant comprising a nucleic acid having or coding for at least one IRES element necessary for translation of a gene required for amplification of said vector and located downstream of said IRES element, said vector further comprising at least a portion of a sequence of the host plant genome in an anti-sense orientation for suppressing a gene of the host plant.

7. Vector according to one of claims 1 to 6 wherein it is derived from a plant virus.
8. Vector according to one of claims 1 to 7 wherein it comprises a gene coding for a protein mediating cell-to-cell or long-distance movement of said vector in a plant.
9. Vector according to one of claims 1 to 8 wherein it codes for protein(s) functional for amplification.
10. Vector according to one of claim 1 to 9 wherein said IRES element is of plant viral origin.
11. Vector according to one of claims 1 to 9 wherein said IRES element is or comprises segment(s) of a natural IRES of plant origin.
12. Vector according to one of claims 1 to 9 wherein said IRES element is a synthetic IRES element.
13. Vector according to claim 12 wherein the IRES element is or comprises a multimer of a segment of a natural IRES element.
14. Vector according to claim 12 wherein the IRES element is or comprises a multimer of at least one sequence essentially complementary to an IRES-binding segment of a natural 18S rRNA.
15. Vector according to one of claims 1 to 14 wherein translation of one or several gene(s) encoded by said vector is cap-independent.
16. Pro-vector having a sequence that is subject to processing by the host plant nucleic acid processing machinery for yielding a vector according to claims 1 to 15.

17. Pro-vector that is convertible *in vitro* or *in vivo* into the vector according to one of claims 1 to 16 by standard procedures of molecular biology.
18. Use of a vector according to one of claims 1 to 15 for determining the function of a structural gene.
19. Use of a vector according to one of claims 1 to 15 for producing a protein.
20. Use according to claim 19, wherein the protein is selected from the group consisting of antibodies, antigens, receptor antagonists, neuropeptides, enzymes, blood factors, Factor VIII, Factor IX, insulin, pro-insulin, somatotropin, serum albumin, tissue-type plasminogen activator, tissue-type plasminogen activator, haematopoietic factors such as granulocyte-macrophage colony stimulating factor, macrophage colony stimulating factor, granulocyte colony stimulating factor, interleukin 3, interleukin 11, thrombopoietin, erythropoietin.
21. Use of a vector according to claims 1 to 15 for generating a trait in the host plant.
22. Use according to one of claims 19 to 21, whereby the vector is applied to plants or parts of plants on a farm field.
23. Use according to one of claims 18 to 22, whereby the plant is treated with an agent inhibiting cap-dependent translation.
24. Gene expression system comprising a vector or pro-vector according to one of claims 1 to 17 and a natural or genetically engineered plant that supports amplification and expression of said vector.

## 59

25. System according to claim 24, further comprising an *Agrobacterium* intermediary host system that supports delivery of one or more of the vectors or pro-vectors according to one of claims 1 to 17 into the plant.
26. System according to claim 25, wherein said *Agrobacterium* intermediary host further supports transfer and transient or stable expression of other traits necessary or desirable for expression of a gene to be expressed.
27. System according to one of claims 25 to 28, wherein said system supports expression of two or more genes in the same plant cell or in the same plant.
28. A system according to one of claims 25 to 29, wherein said gene to be expressed is selected from the group consisting of antibodies, antigens, receptor antagonists, neuropeptides, enzymes, blood factors, Factor VIII, Factor IX, insulin, pro-insulin, somatotropin, serum albumin, tissue-type plasminogen activator, tissue-type plasminogen activator, haematopoietic factors such as granulocyte-macrophage colony stimulating factor, macrophage colony stimulating factor, granulocyte colony stimulating factor, interleukin 3, interleukin 11, thrombopoietin, erythropoietin.
30. A method for generating a vector according to one of claims 1 to 17, whereby the IRES element is produced through directed evolution from a randomized nucleotide sequence by  
  
selecting IRES elements necessary for translation of a reporter gene downstream thereof.



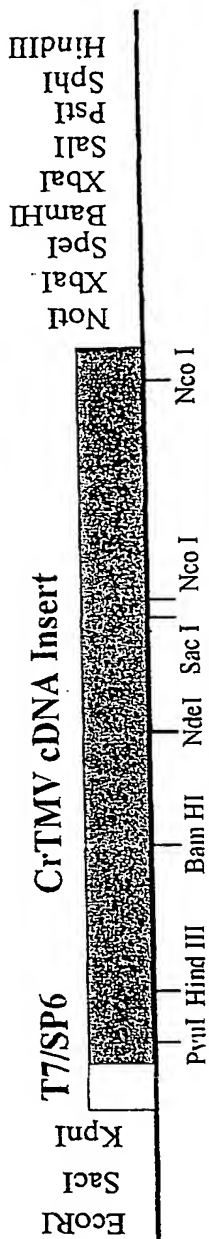


Fig.1

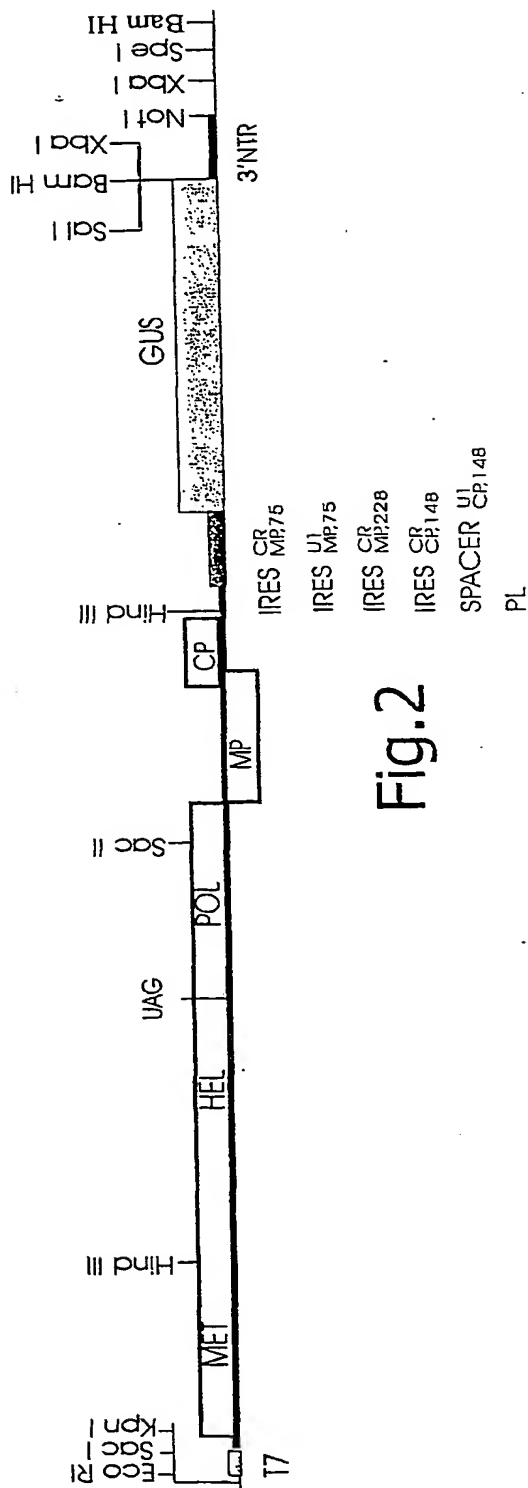
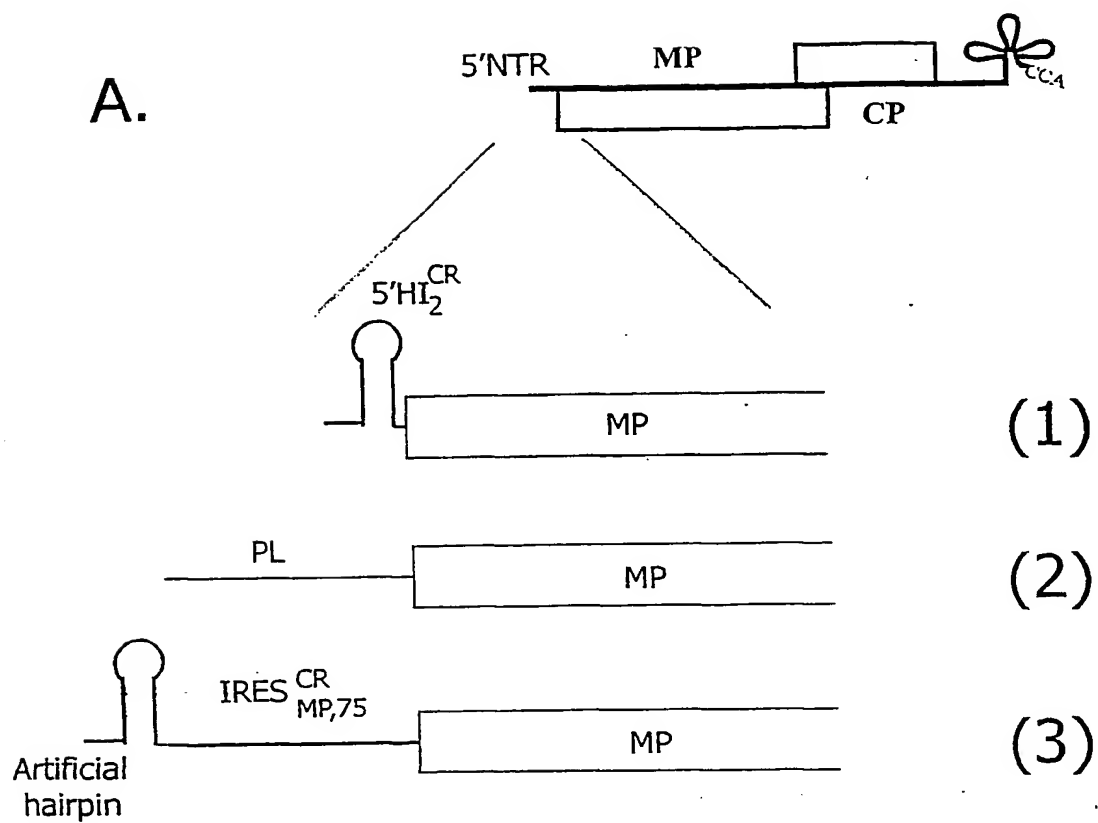


Fig.2





B. TRANSLATION OF  
TRANSCRIPTS 1, 2, 3 IN WGE

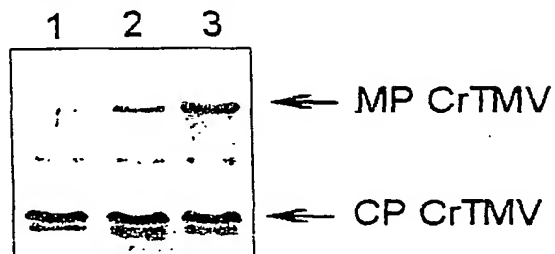


Fig.4

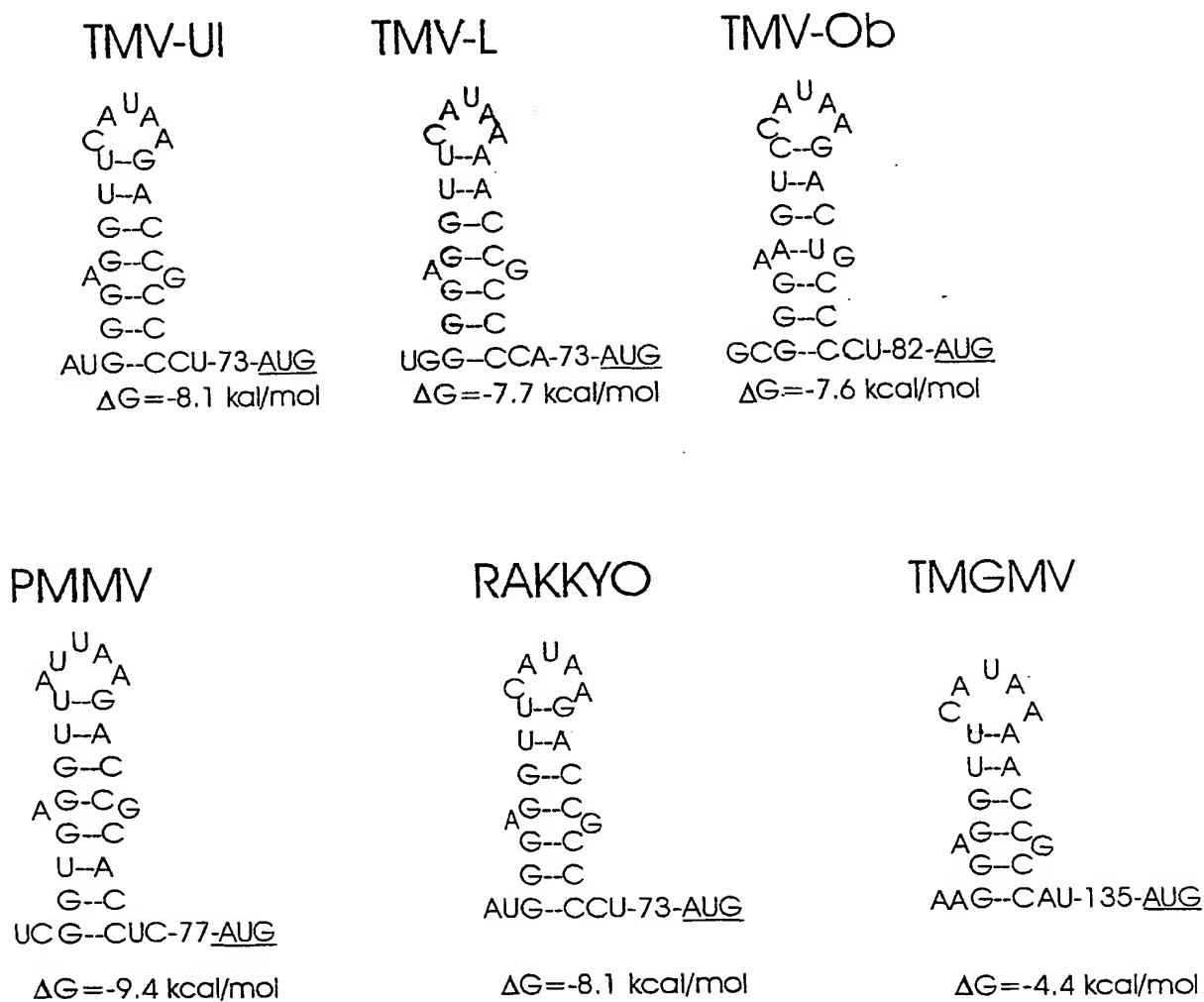


Fig.5

- A, B. RNA-tag with known sequence is ligated specifically to the cap of tested RNA.
- C. Reverse transcription with 3'-specific primer and synthesis of first-strand cDNA. Tag sequence is included to the sequence of cDNA.
- D. PCR with tag-specific and 3'-specific primers. The appearance of the respective PCR band indicates the presence of cap-structure in the tested RNA.
- E. PCR with 5'-specific and 3'-specific primers. The appearance of PCR band serves as control for PCR reaction and indicates a presence of the specific tested RNA in the reaction.
- F. Relative comparison of the lengths of obtained PCR bands.

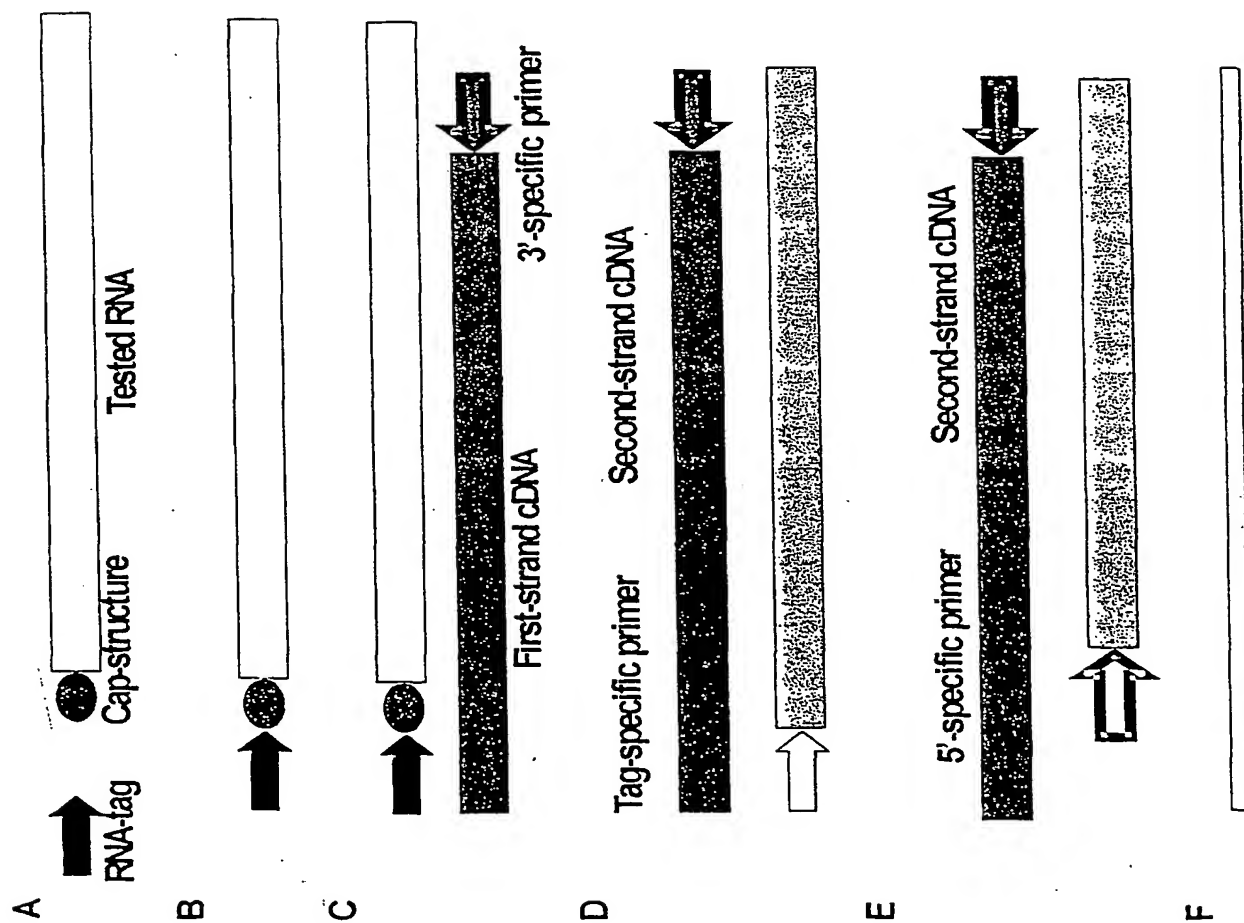


Fig.6

PCR products with 3'-specific/tag-specific and 3'-specific/5'-specific primers

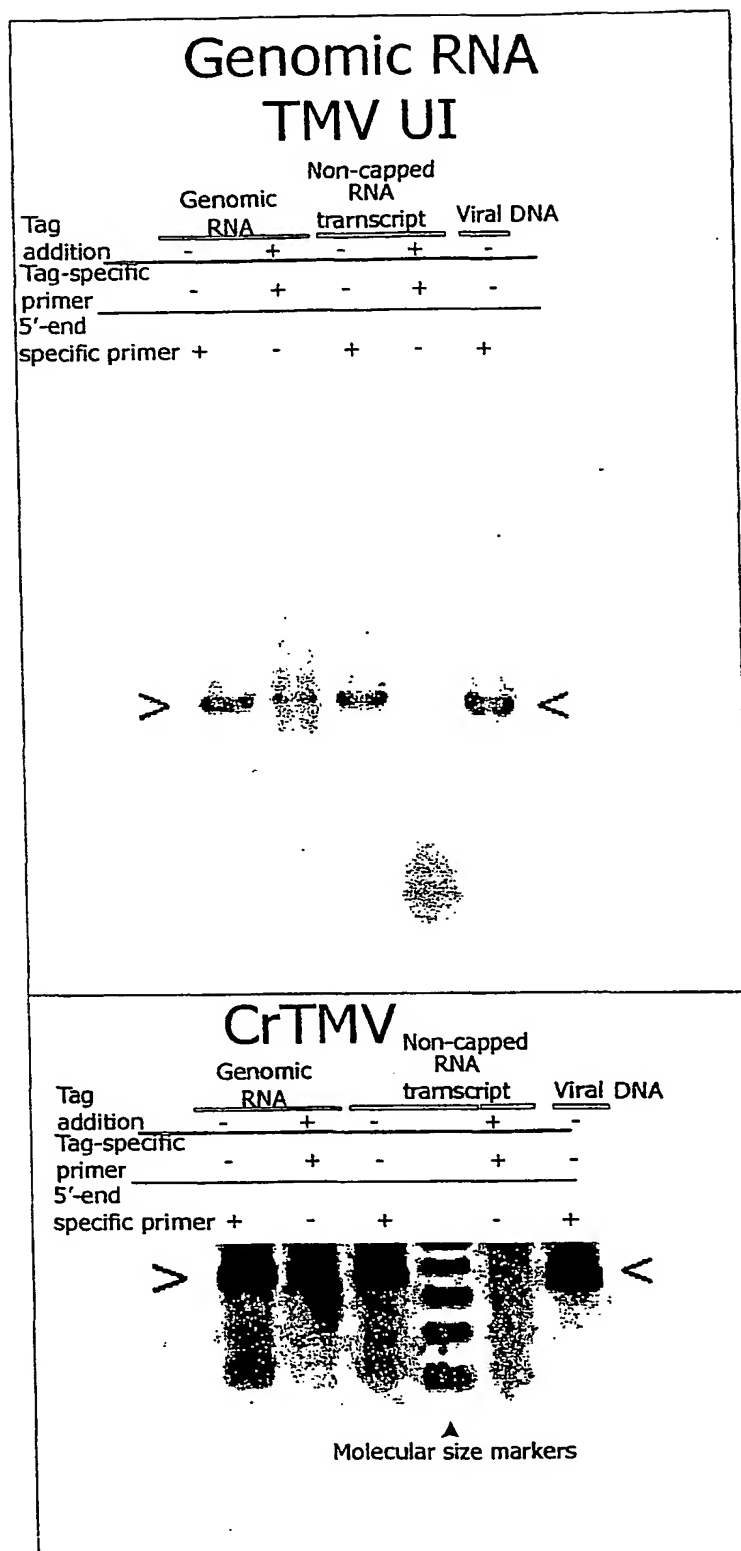
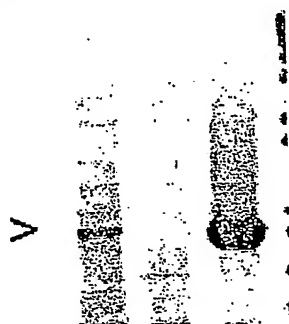


Fig.7a

## I<sub>2</sub> subgenomic RNA TMV UI

	I <sub>2</sub> RNA		Viral DNA
Tag addition	-	+	-
Tag-specific primer	-	+	-
5'-end specific primer	+	-	+



## CrTMV

	I <sub>2</sub> RNA		Viral DNA
Tag addition	-	+	-
Tag-specific primer	-	+	-
5'-end specific primer	+	-	+



Fig.7b

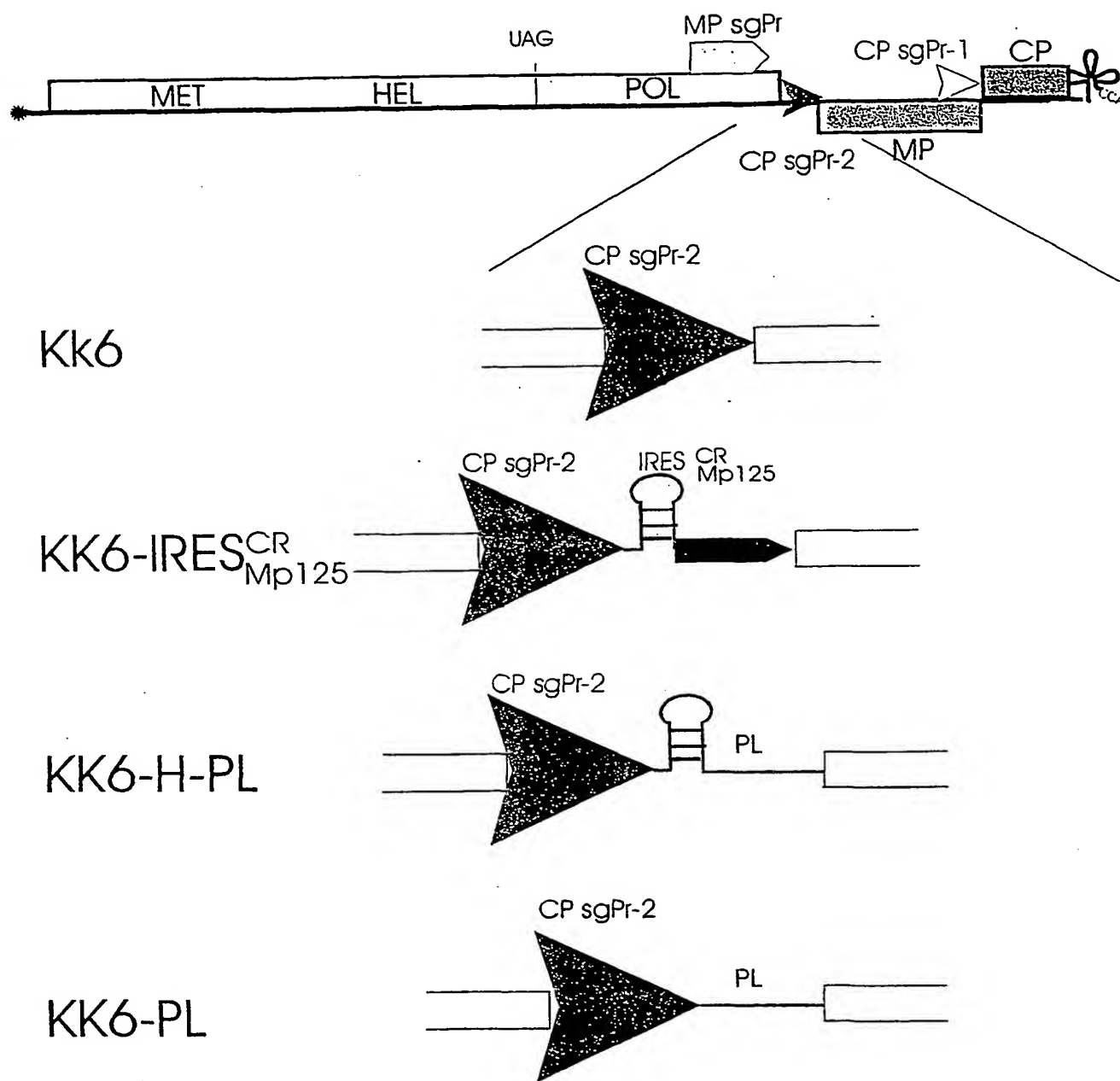


Fig.8



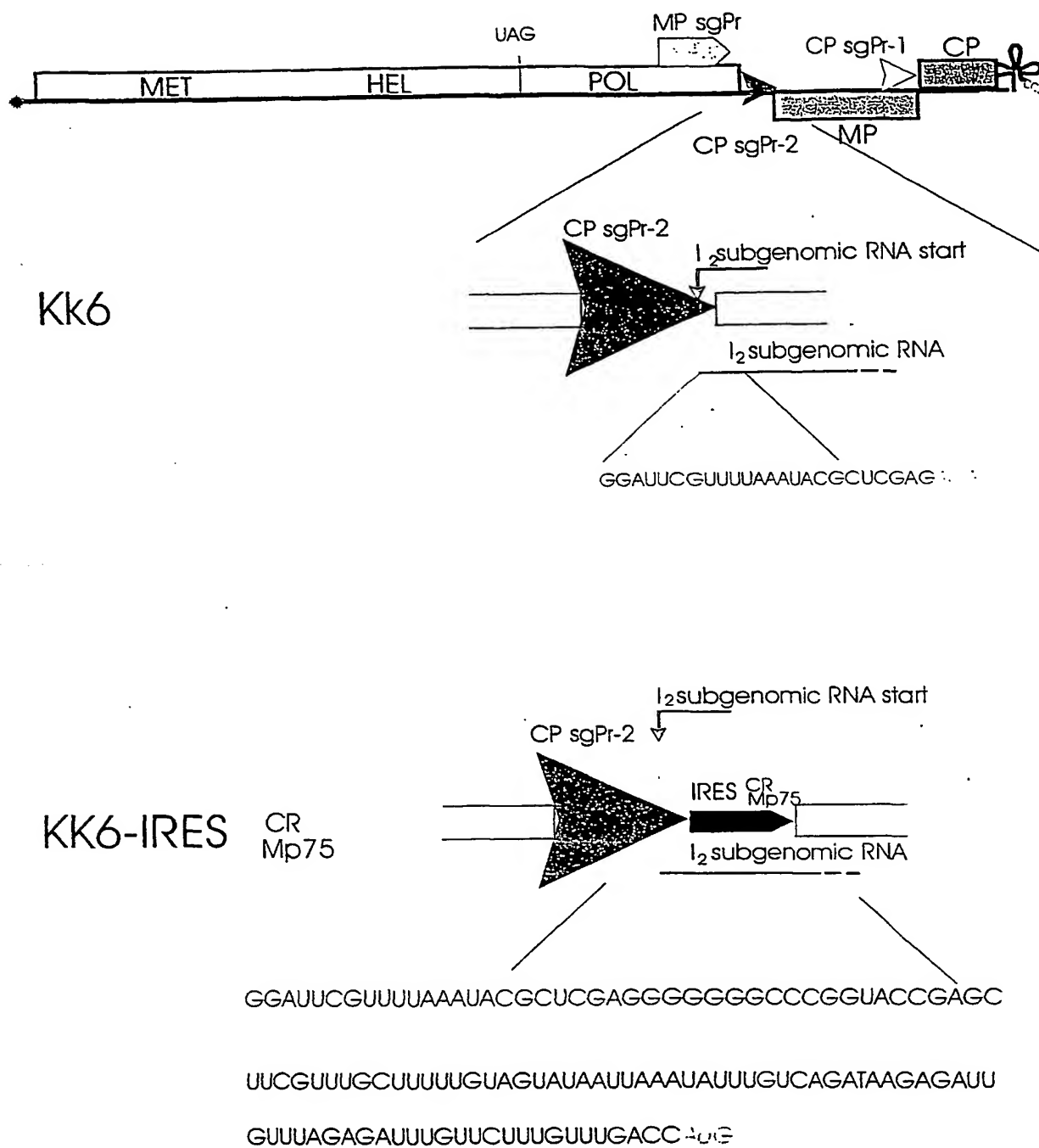
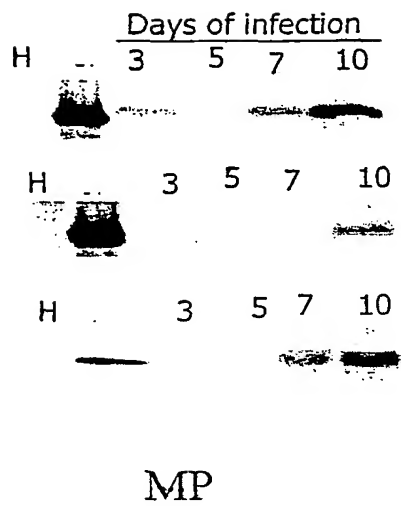


Fig.9



UI

Kk6

K86

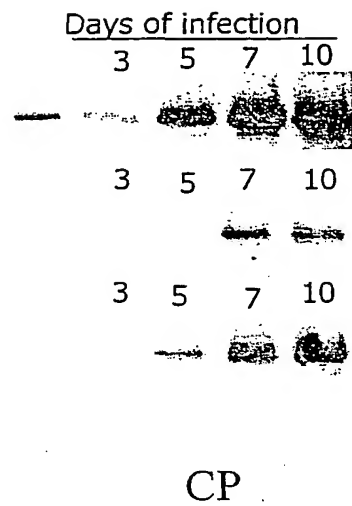
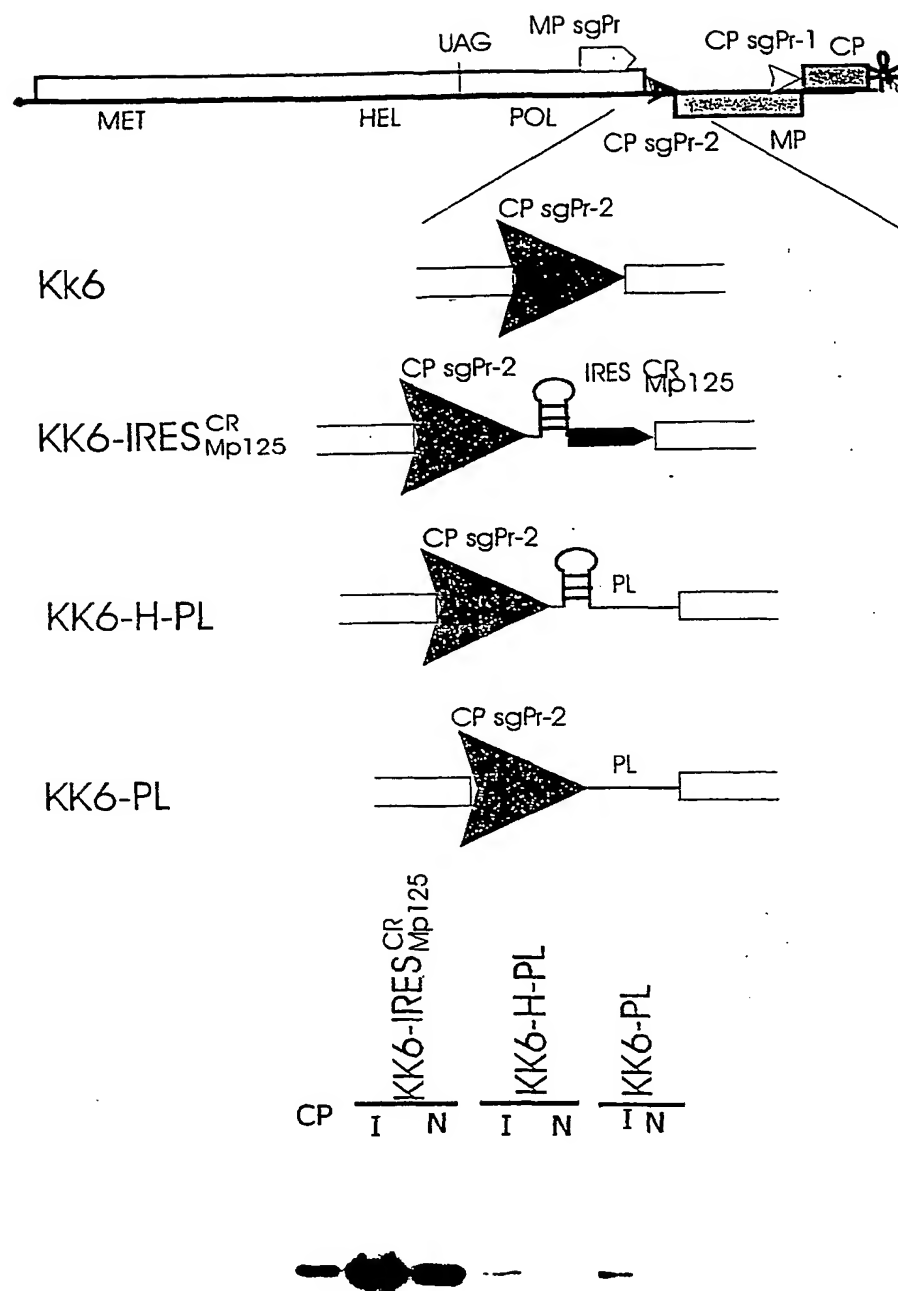


Fig.10



I- Inoculated leaves  
N- noninoculated upper leaves

Fig.11

CrTMV IRESmp multimer structure  
and complementarity to *A. thaliana*  
18S rRNA

A.th. 18S rRNA 779-3'-CGAACGAAACUC-5'-761  
IRES -75-5'-UUCGUUUGCUUUUUGUAGUAUAUUAUUUGUCAGAUAA-25  
-34-GAGAUUGUUUAGA-22  
-21-GAUUUUGUUC-13  
-12-UUUGUUUGAUAAUG-+3

Fig.12

13/30

36nt-Polylinker

— HGFP-PL-GUS

— HGFP-IRES<sub>MP,75</sub>-GUS

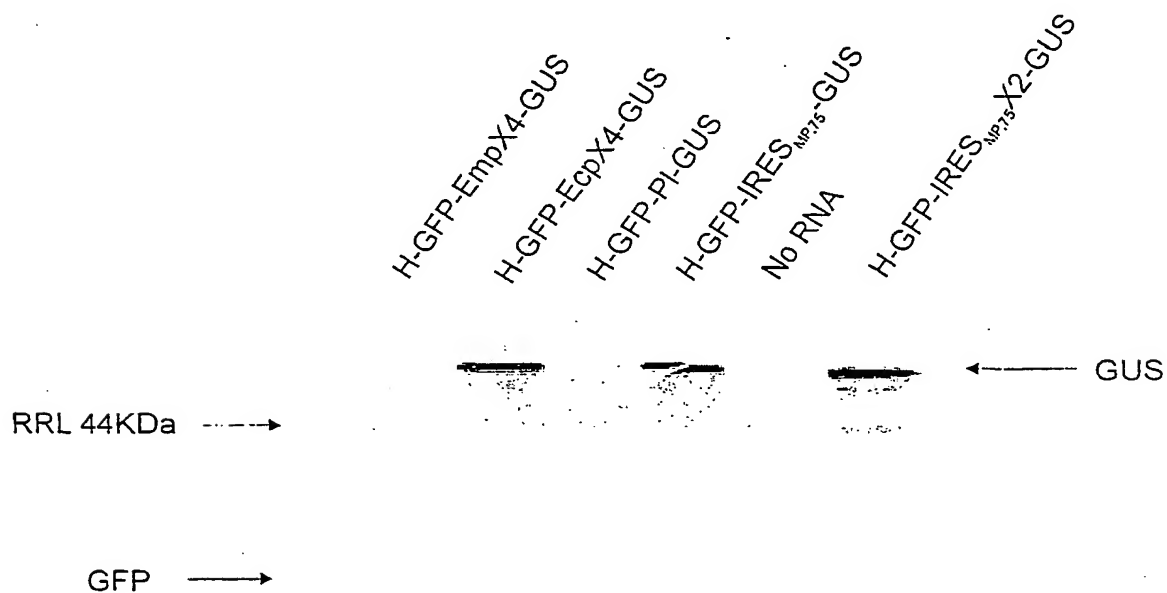
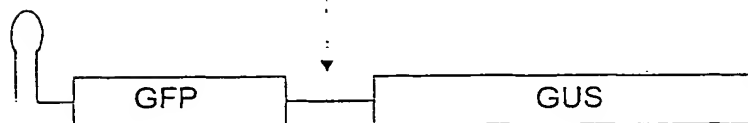
IRES<sub>MP,75</sub><sup>C<sub>1</sub>TMV</sup>

IRES<sub>MP,75</sub><sup>C<sub>1</sub>TMV</sup> + IRES<sub>MP,75</sub><sup>C<sub>1</sub>TMV</sup>

— HGFP-IRES<sub>MP,75</sub>-GUS

(CGUUUGCUUUUUGUAGUA) X 4 — HGFP-EmpX4-GUS

(AAAAGAAGGAAAAAGAAGG) X 4 — HGFP-EcpX4-GUS



*In vitro* Translation in RRL

Fig.13

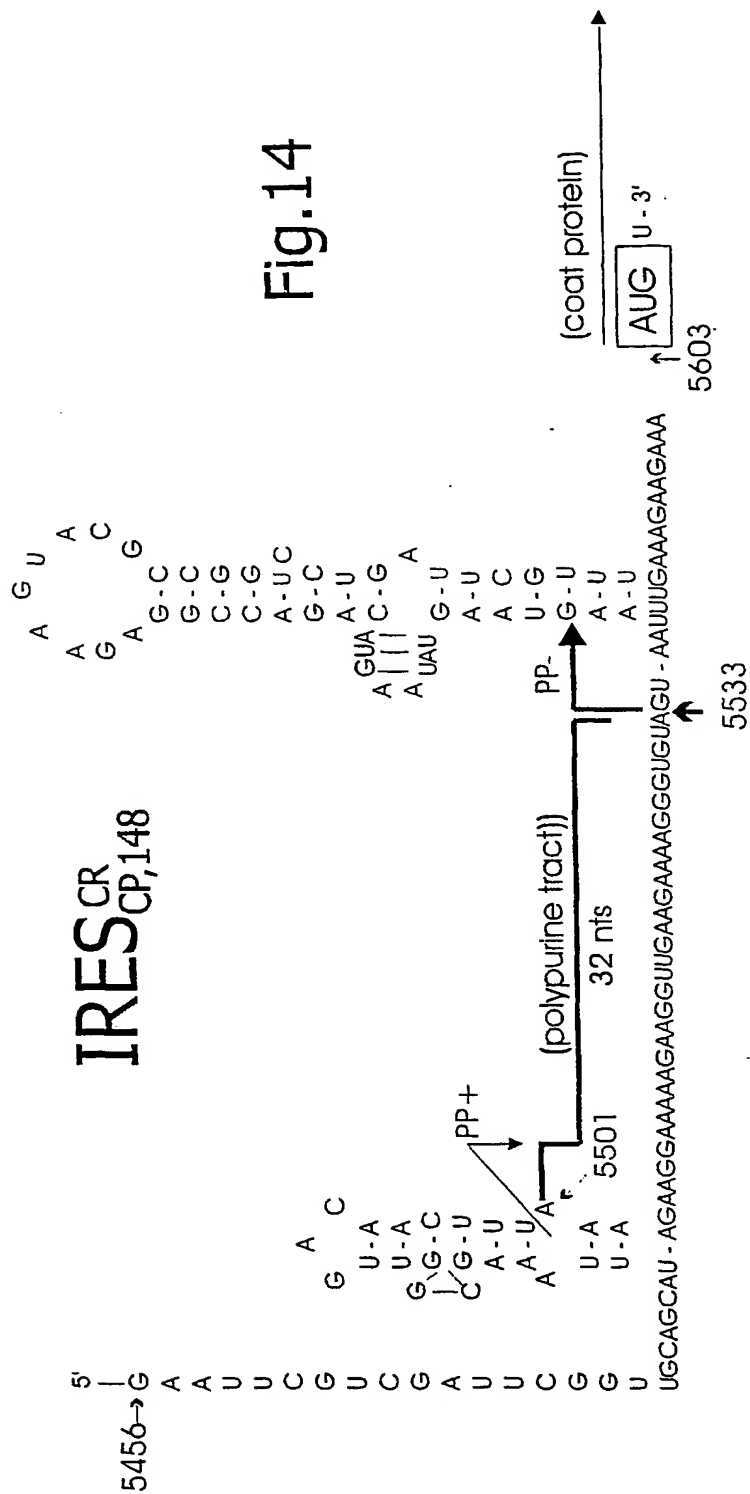
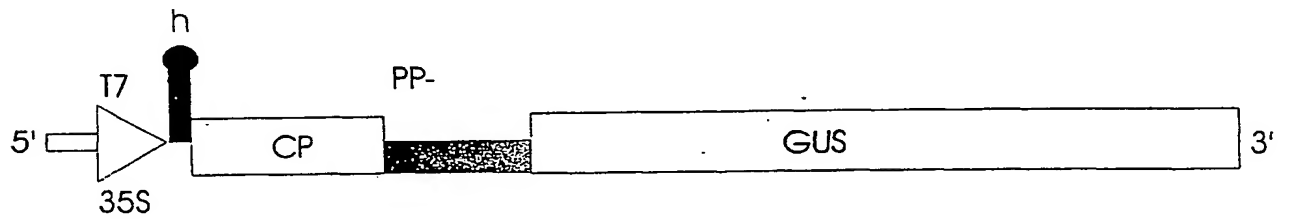
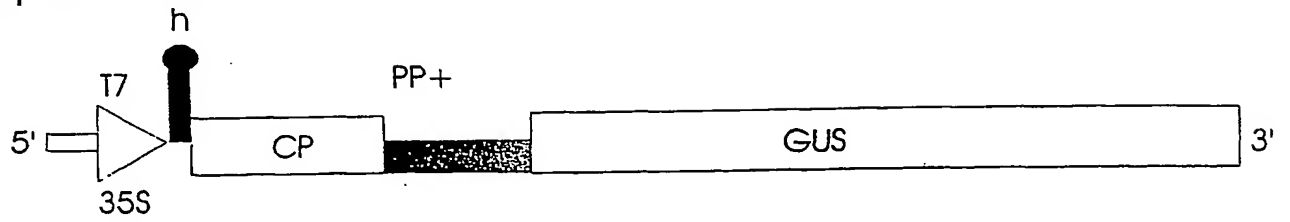
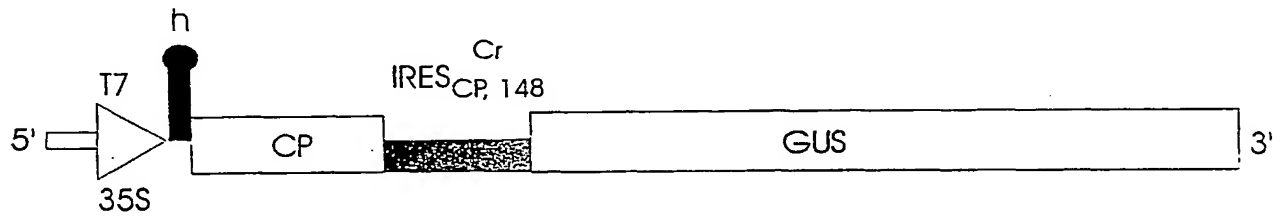


Fig.14

## Experimental constructs



## Positive control



## Negative control

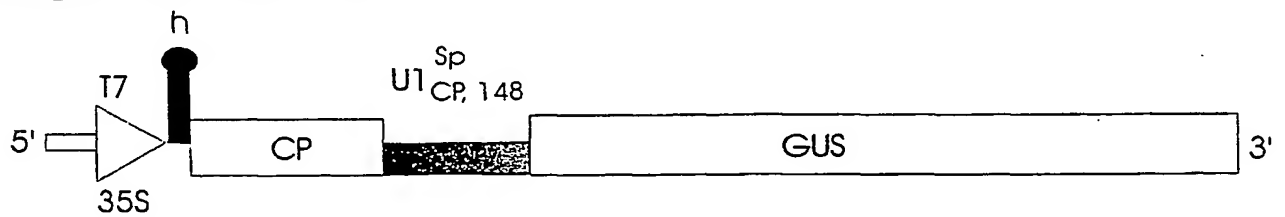


Fig.15

PP+ and PP- deletion mutants  
translation in wheat germ extract

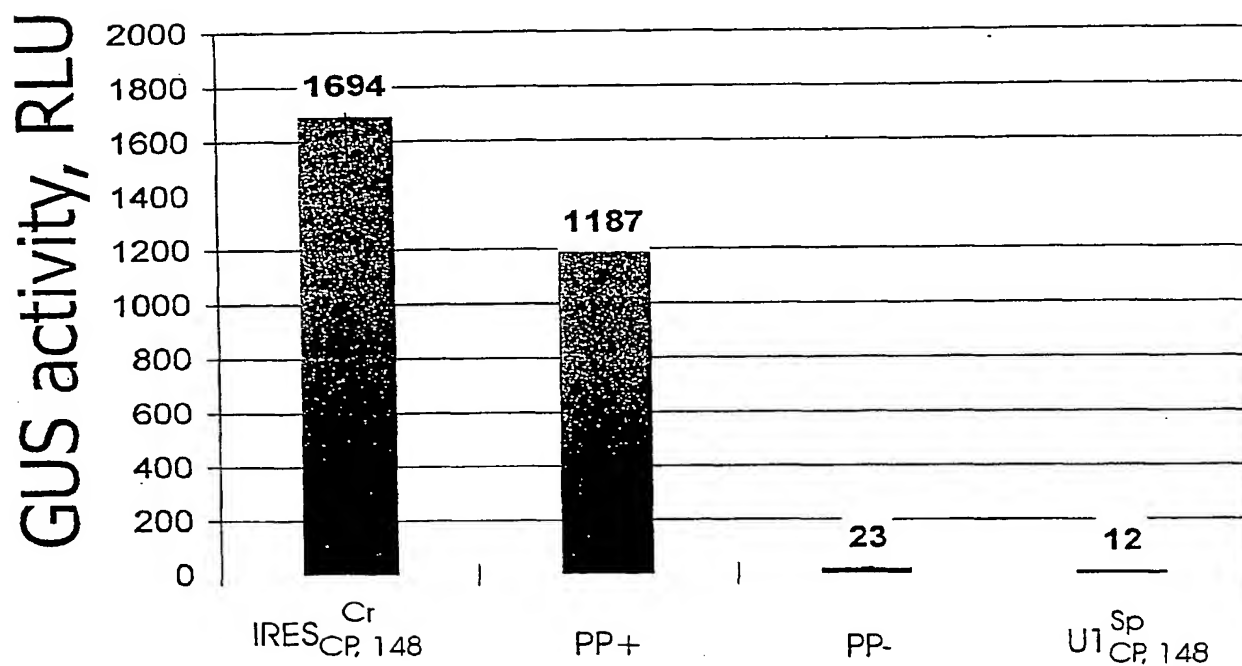


Fig.16



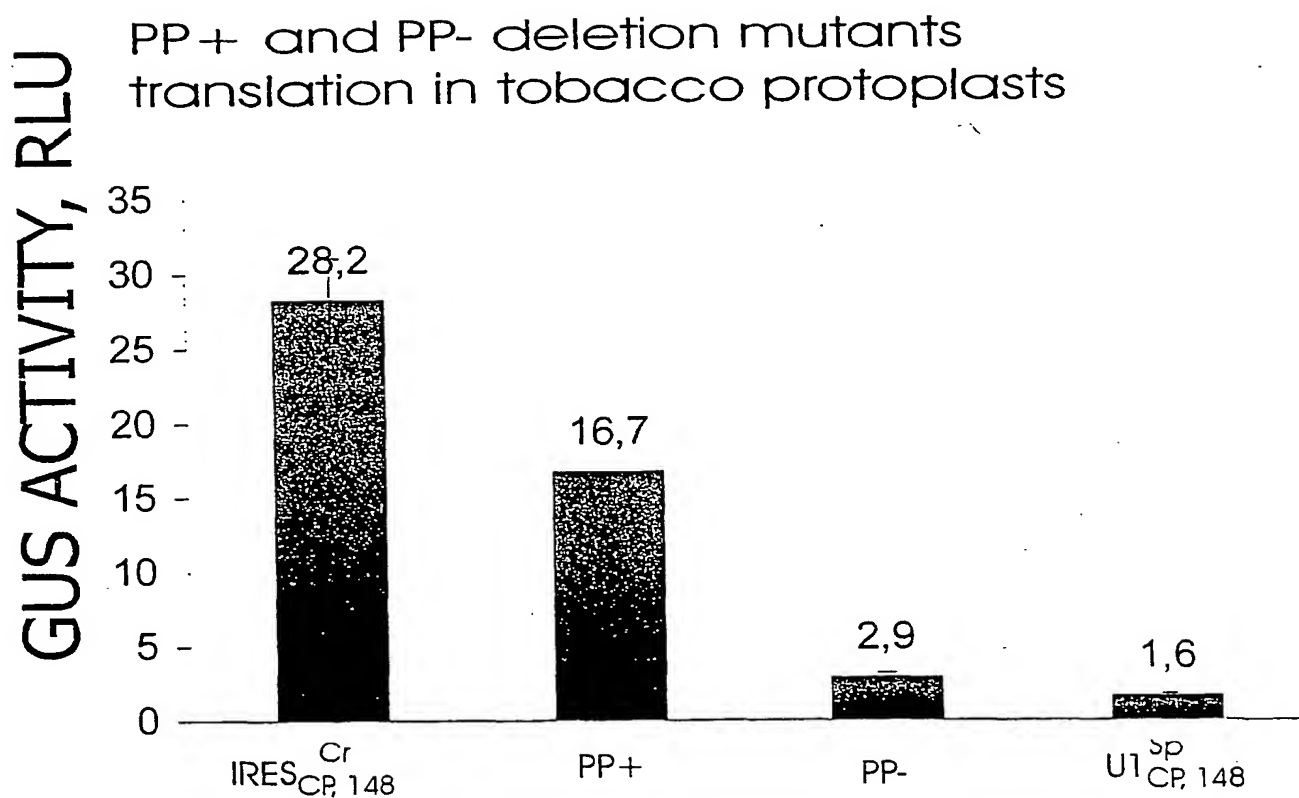


Fig.17

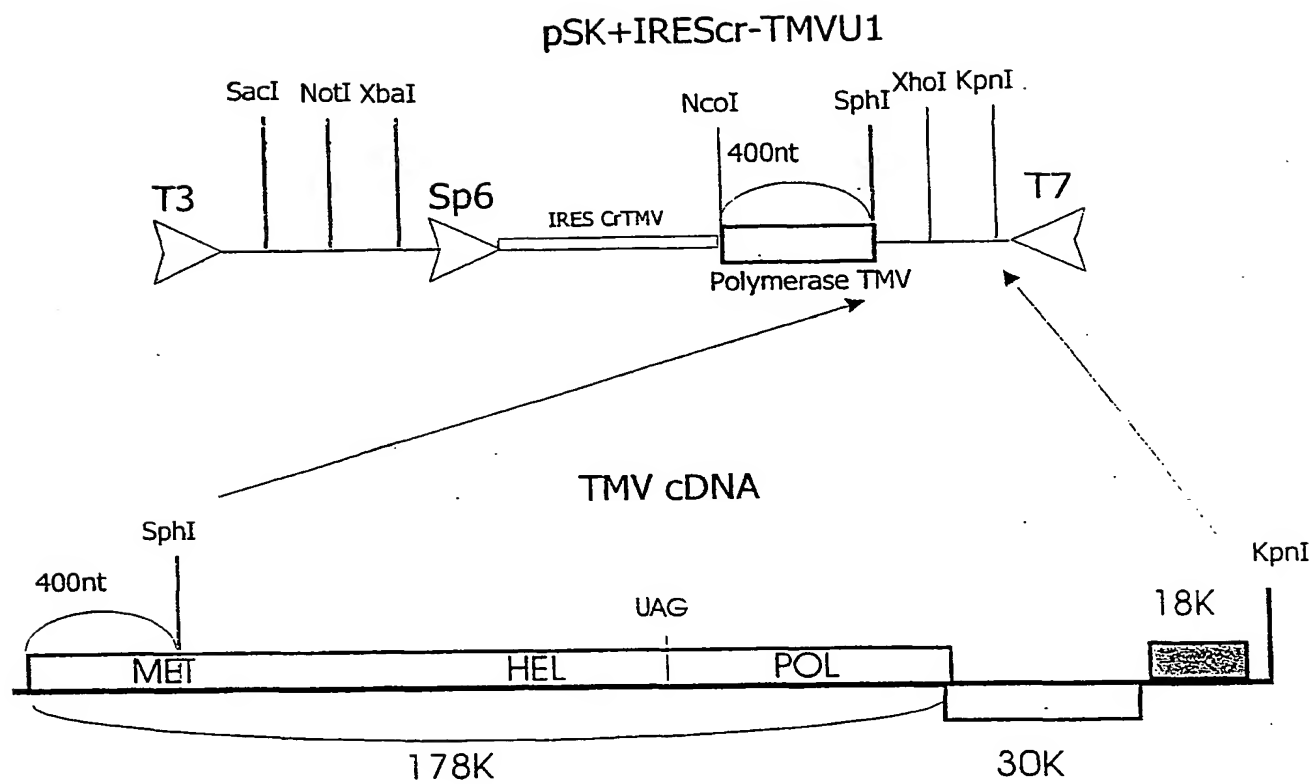


Fig.18

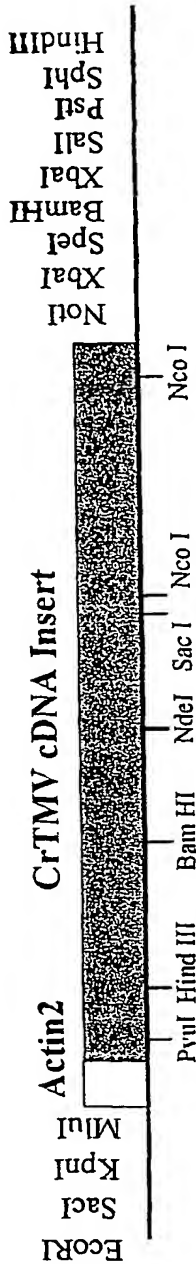


Fig.19

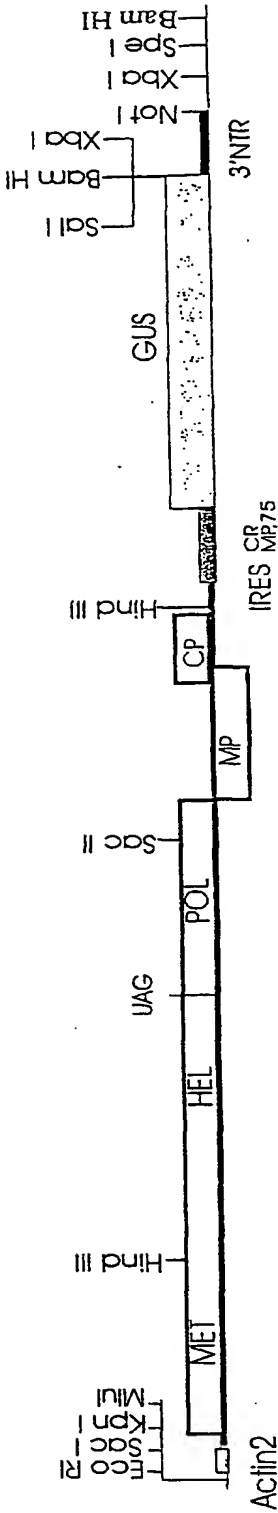


Fig.20

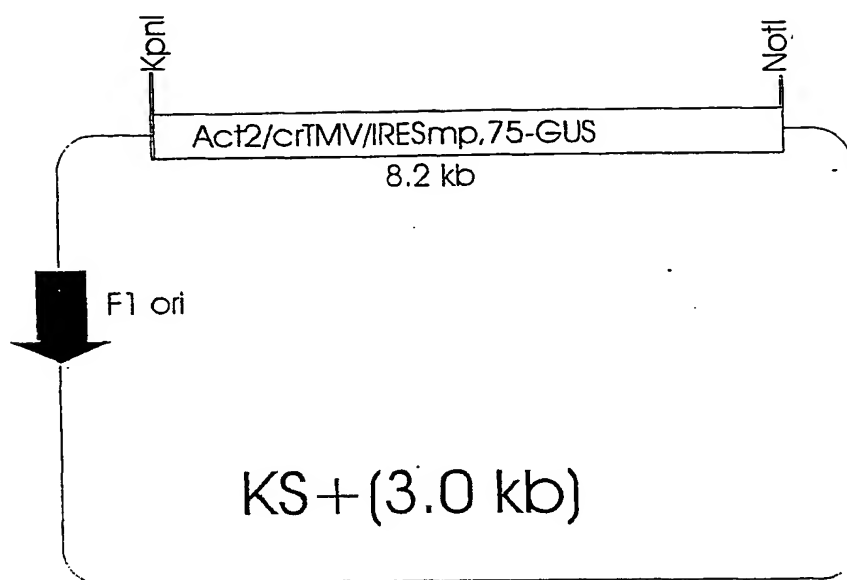


Fig.21

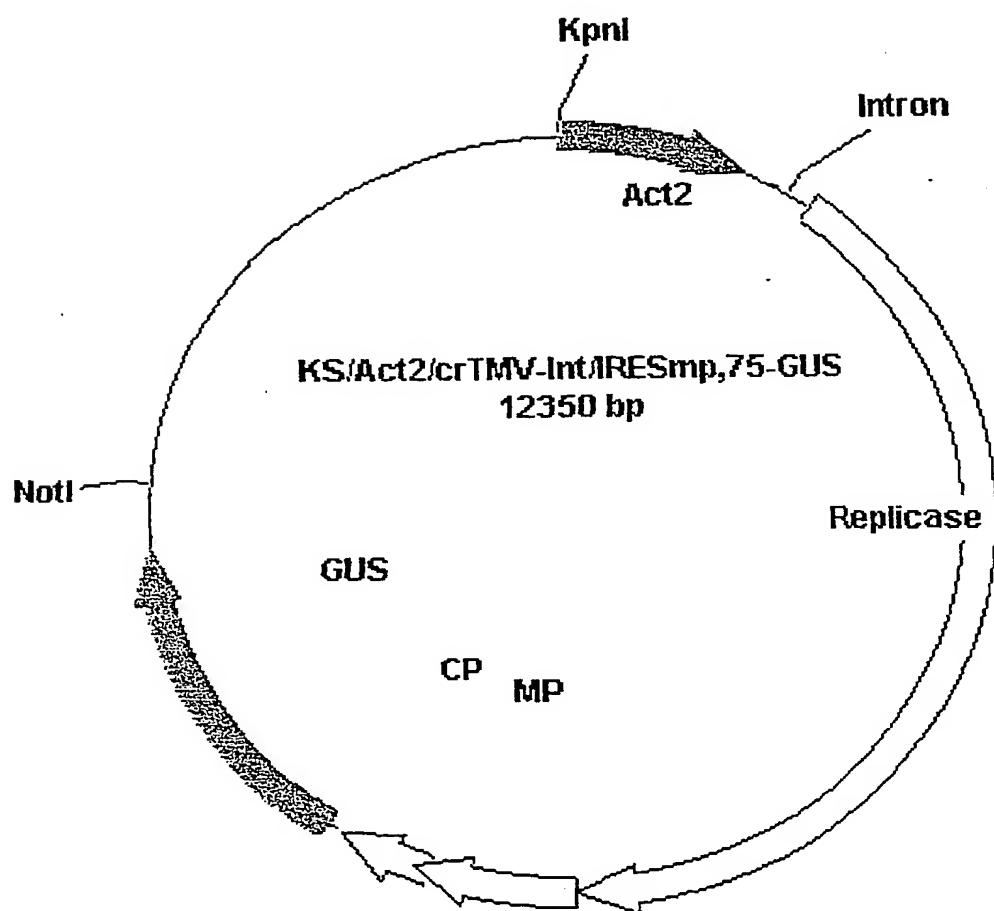


Fig.22

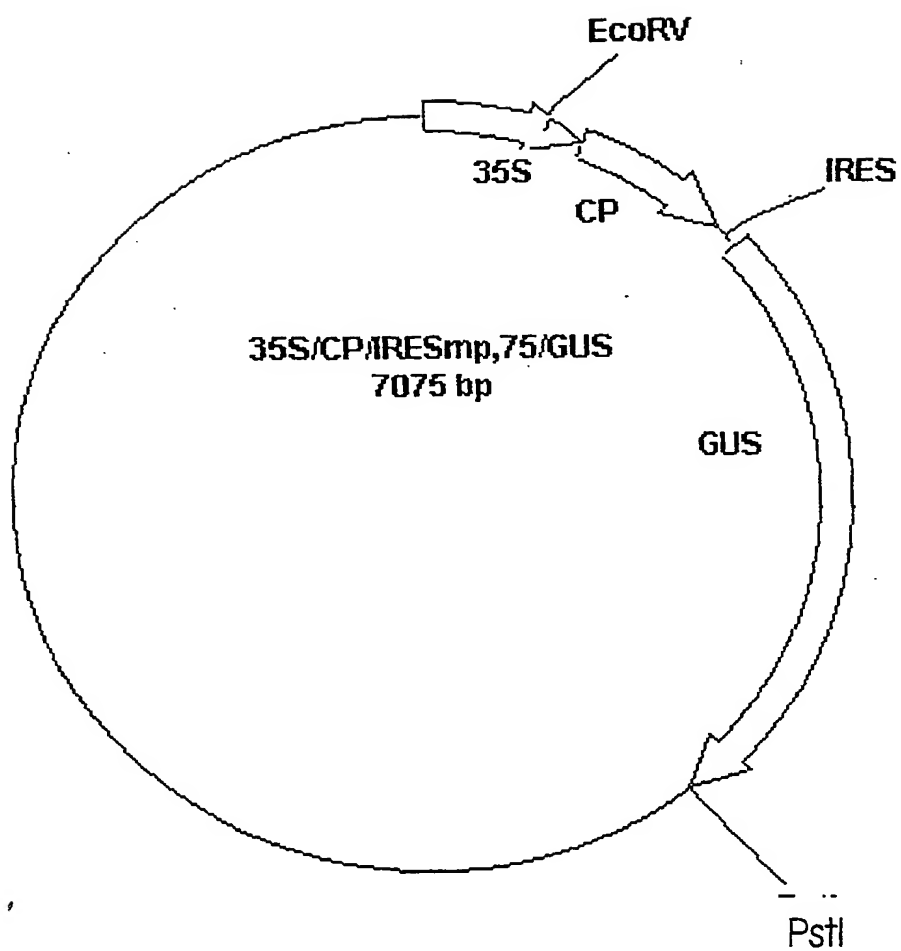


Fig.23

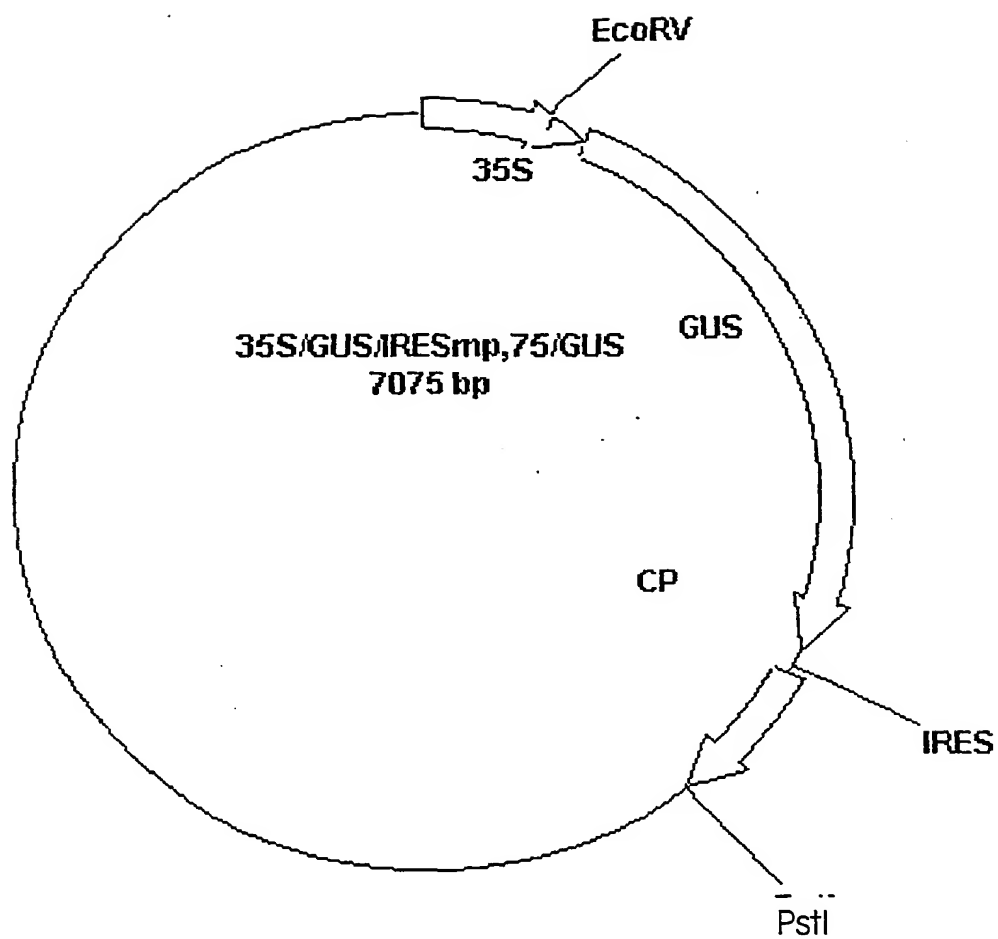


Fig.24

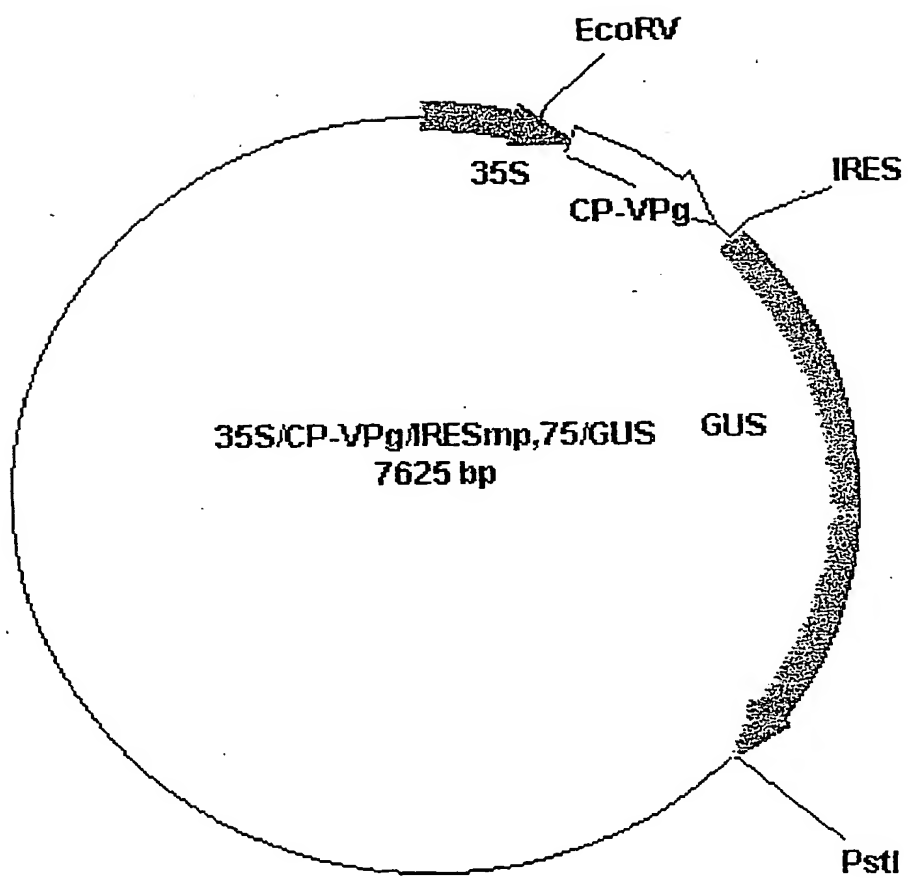


Fig.25



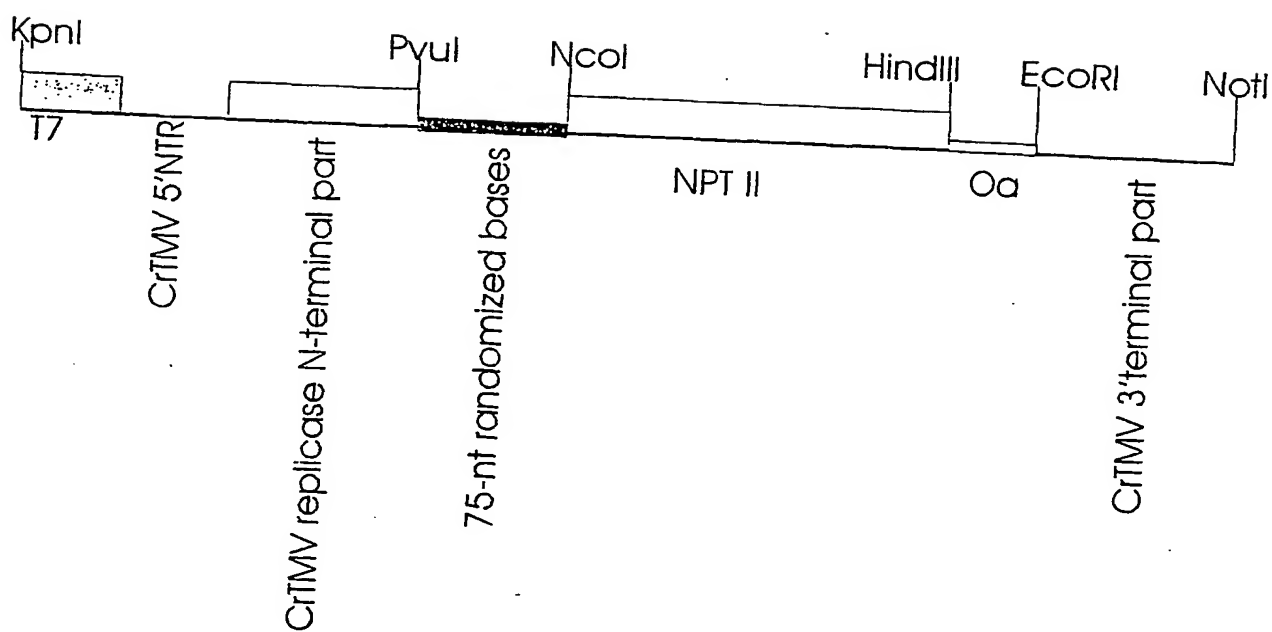
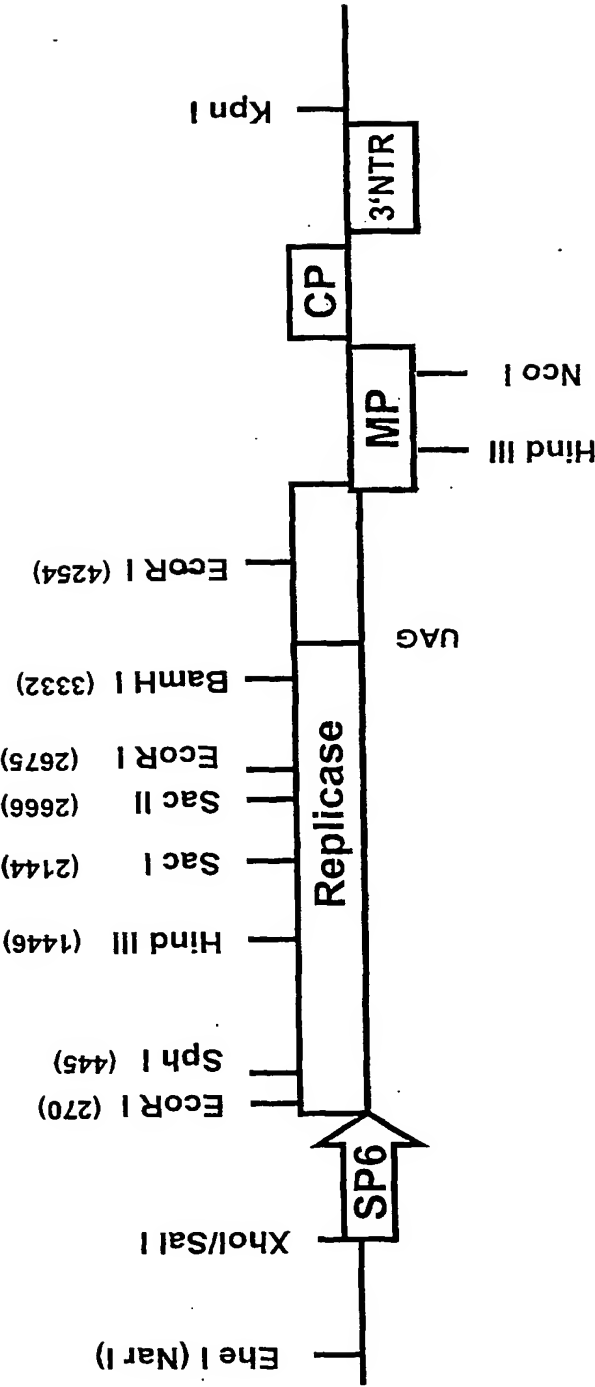


Fig.26

TMV304



Kpn I – for linearization

pBR322 – based plasmid vector,  
Ampicillin resistance

Fig.27  
3'-NTR = nontranslated region  
(Pseudoknots and t-RNA-like  
Structure)

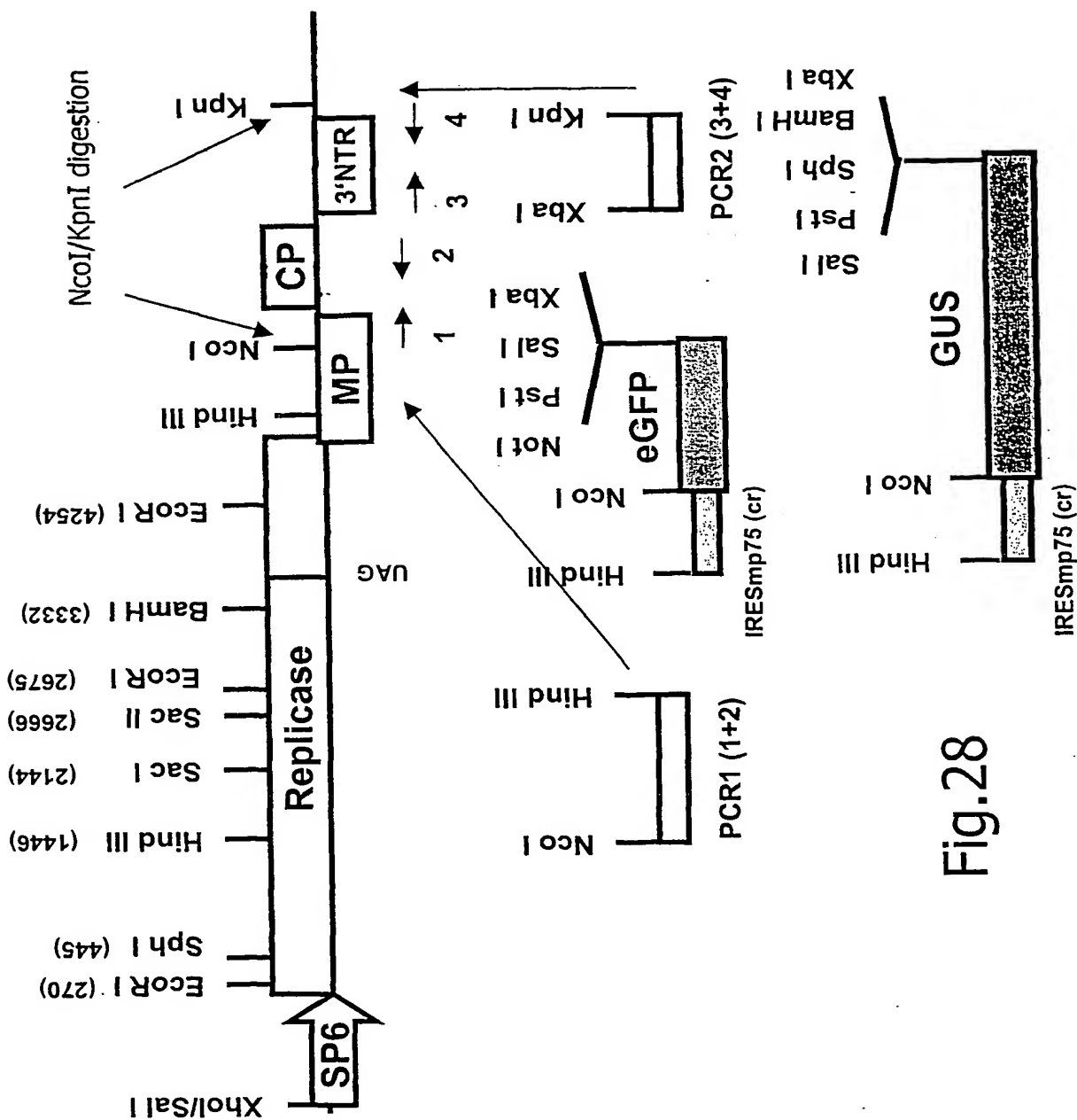
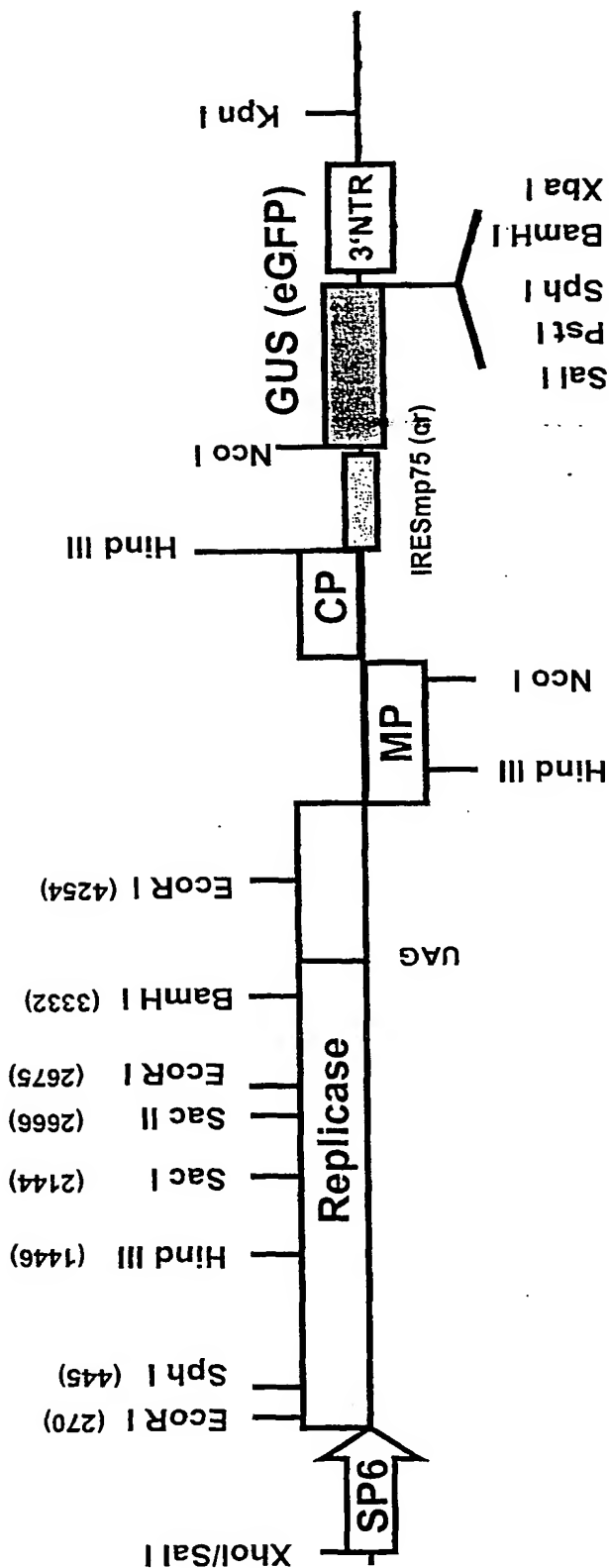


Fig.28

# TMVU1- GUS (GFP)



Kpn I – for linearization

pBR322 – based plasmid vector,  
Ampicillin resistance

3'-NTR = nontranslated region  
(Pseudoknots and t-RNA-like  
Structure)

eGFP – enhanced (red shifted) mutant  
Excitation 488 nm, emission 510 nm

Fig.29

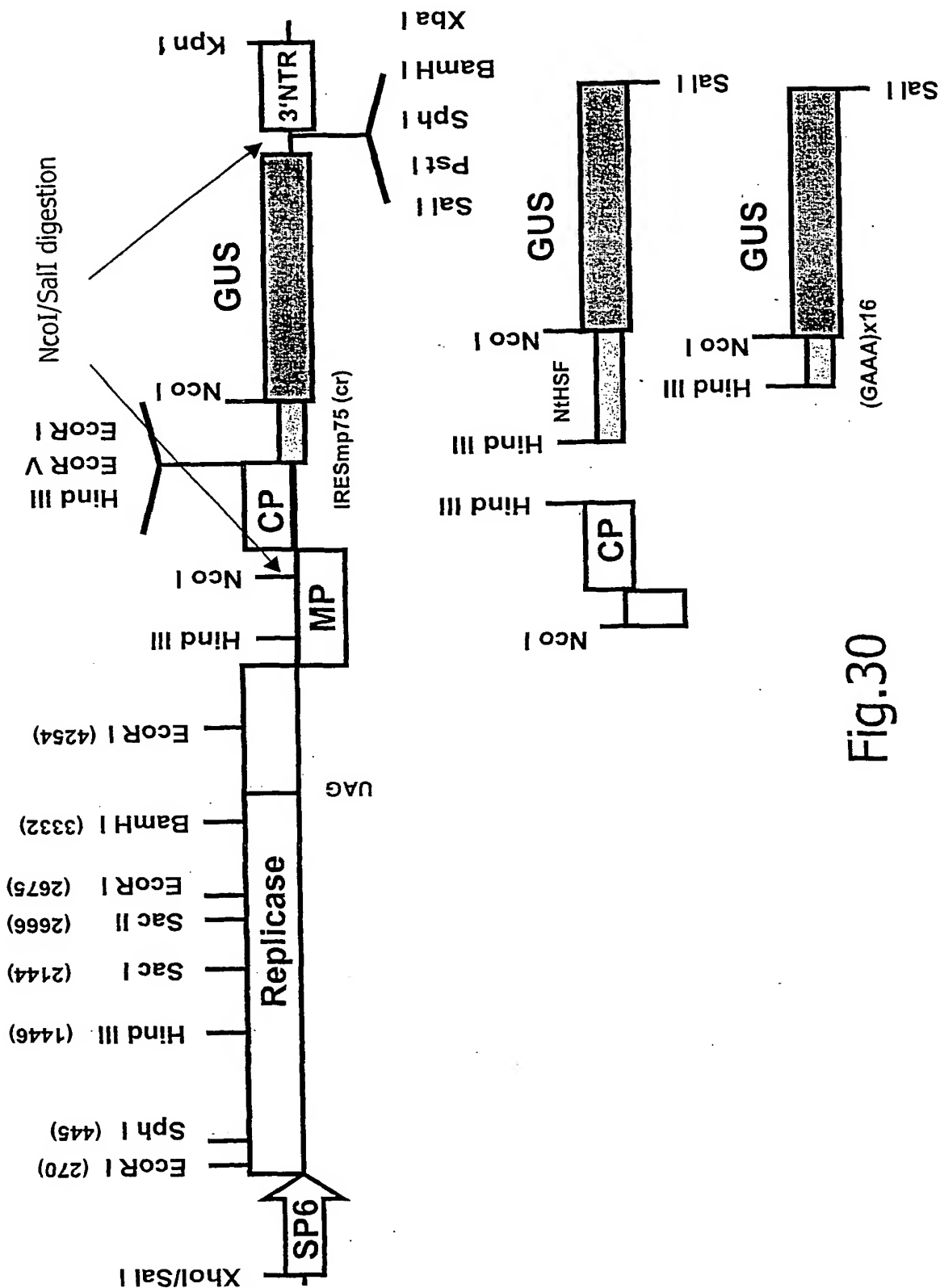
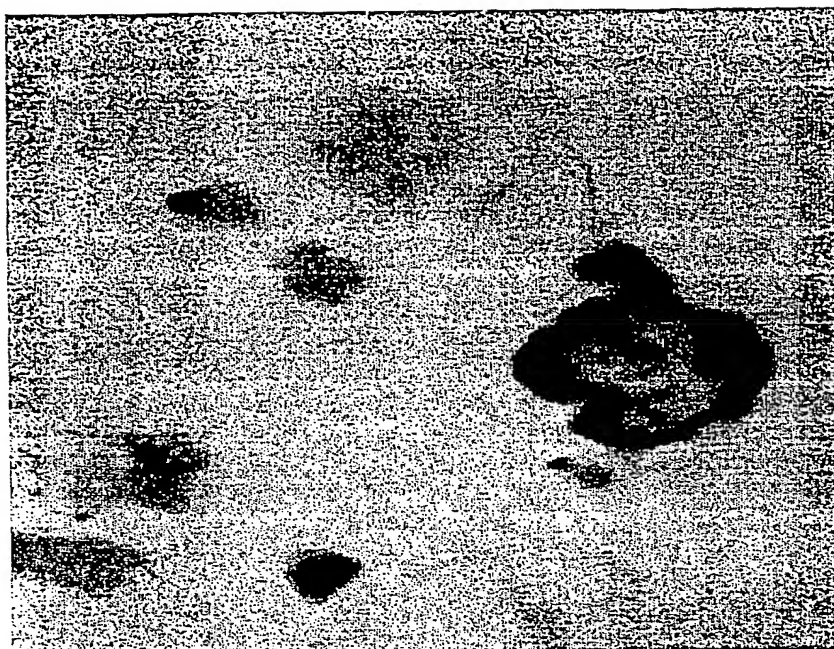


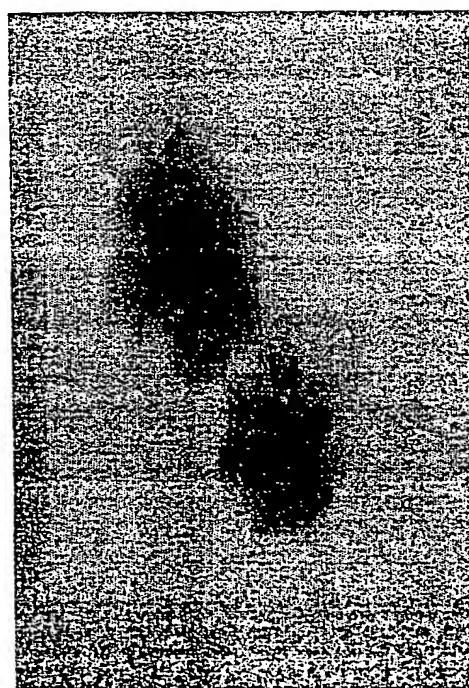
Fig.30



C



A



B

Fig.31

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ ~~FADED TEXT OR DRAWING~~
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**